ROTIS ON THE RARA:
LEXICAL NATIVISATION IN FIJI ENGLISH

Jan Tent: *Department of Linguistics, Macquarie University*
(Sydney, NSW 2109, Australia) <jtent@ling.mq.edu.au>

Abstract

In outer circle Englishes, the boundaries between nativised loanwords and code-switched lexical items are often blurred. This article attempts to gauge the degree of nativisation of a set of common Fijian and Hindi loans by surveying expatriates' use and knowledge of these words. Results indicate that loans forming part of expatriates' active vocabularies also occur most often in the media and in conversation. These words also are pragmatically essential for effective day-to-day communication in Fiji. Differences in the types of words used and known by women, men and teenagers also reflect the immediate pragmatic value they have for each of these groups. The number of loans expatriates acquire increases steadily over the first seven years of residence. After this few new Fijian or Hindi words are learned.

1. Introduction

Fiji belongs to Kachru's 'outer circle' of English speaking nations. Although English is one of the three major languages spoken in Fiji (the other two being Fijian and Fiji Hindi[1]), it is the first language of only a tiny portion of the population (1-2%). Nevertheless, it plays a pivotal role in the day-to-day lives of most, if not all, Fiji Islanders. English is the principal language of government, administration, the judicial system, and commerce; the major, and sometimes the only, medium of instruction in the education system; and an important, though by no means the *only*, lingua franca among people with different first languages. English is also the dominant language of the media.

The structure, development and nativisation of the Fiji English lexis is much the same as that of any other variety of post-colonial English. What sets it apart from all other varieties of English is the rich and colourful amalgam of Fijian and Hindi expressions. Hindi words found their way into Fiji English (as well as Fijian) after the introduction of large numbers of Indian indentured labourers between 1879 and 1916.

During my seven year residence in Fiji I compiled a corpus of Fiji English lexemes and expressions. The corpus is based on numerous written and oral sources. These include: stories, articles, letters, and advertisements in the local English print media; university students' essays, assignments and examination scripts; hand-written and printed notices and signs; locally published plays and novels; excerpts from conversations I either overheard or personally participated in; recorded interviews and conversations; as well as television and radio news broadcasts, commentaries, advertisements, and community announcements.

The lexis of Fiji English embodies much of the same type of lexical material of other post-colonial Englishes. It comprises: 'indigenous' loans (i.e. Fijian, Hindi, and Polynesian); loans from other varieties of English (e.g. Indian, Australian, American and British English); calques (mainly from Fijian); reborrowings[2]; hybrids (i.e. English + Fijian or Hindi lexical collocations and compounds); standard English lexemes that have undergone locally motivated semantic shifts; grammatical conversions; novel compounds of existing standard English lexemes; English archaisms; neologisms; and locally coined exclamations, interjections and directives.

The majority of Fiji English speakers are at least bilingual, generally having as their first language either Fijian or Fiji Hindi.[3] The greater part of indigenous Fijians and Indo-Fijians are also quite conversant in each other's languages. The result is widespread code-switching accompanied by the systematic and routine insertion of Fijian and Hindi words into English. Since interlocutors are often so familiar with each others' languages, the boundaries between nativised loanwords and code-switched Fijian/Hindi lexical material becomes blurred. Naturally, this is not an issue for any bi- or multilingual speaker. It is only an issue for the linguist or lexicographer who feels obliged to delineate the lexical parameters of language varieties.

A continuum of relationships exists between borrowings and all code-switched forms so that the two are not, I believe, clearly distinct phenomena as some (e.g. Poplack 1980; Sankoff, Poplack & Vanniarajan 1990) have suggested. Moreover, the process of lexical nativisation also forms a

continuum, and the point at which an item can or cannot be said to be fully integrated into the recipient language is, therefore, arbitrary.

2. The Survey

When two expatriate Australians invited me to join them at lunchtime for 'rotis on the rara'[4], I came upon the idea of devising a method of determining which Fijian and Hindi words had filtered through into the English of expatriates living in Fiji. This might provide some clue or indication as to which Fijian and Hindi words had become fully nativised loans.

Generally, expatriates have little incentive to learn Fijian or Hindi, especially in urban Suva, where English is used more than almost anywhere else in the country. Furthermore, expatriates are usually under contract and most leave Fiji after completing one or two three-year contracts. Fiji Islanders will invariably address expatriates in English because (a) the chance that an expatriate can speak any Fijian or Hindi is negligible, and (b) Fiji Islanders know expatriates have little or no interest in learning Fijian or Hindi.

Apart from some discourse particles (e.g. *bula* 'hello', *moce* 'good-bye' and *acha* 'okay, fine, good') and exclamations (e.g. *oilei* 'an expression of surprise' and *uro* 'an exclamatory approbation')[5], the most common direct exposure expatriates have to most Fijian and Hindi words is through the print media. Indeed, Deverson (1984: 5) maintains that the principal ways in which Māori has impacted upon New Zealand English is through the print media and literature. In the Fiji context, however, literature cannot be counted as a significant vector for the introduction of Fijian or Hindi loans into Fiji English, as there is an extreme paucity of local literature.

Table 1 shows the typical categories (and their proportions) of Fijian and Hindi loans in Fiji English. The data are based on an analysis of a corpus of more than 680 attested lexical items.

2.1 Method

I devised a simple self-reporting survey similar to that used by Bellett (1995) in her study of Māori lexical influence on New Zealand English. Respondents were presented with a list of 98 Fijian and 45 Hindi words arranged in random order in three columns. The words chosen represented those that have been regularly cited as belonging to Fiji English (see for instance: Siegel 1987, 1989; Monsell-Davis 1984; Arms 1975; Kelly 1975; Moag & Moag 1977;

CATEGORY OF LOAN	% FIJIAN	% HINDI
A. Inedible flora (trees and medicinal herbs)	10	—
B. Edible flora (vegetables, fruits, herbs and spices)	10	12
C. Inedible fauna (birds)	1	—
D. Edible fauna (mainly seafood)	7	1
E. Prepared foods and beverages	3	30
F. General/miscellaneous nouns, adjectives and verbs	27	32
G. Culture and religion (customs, concepts, events and practices)	20	21
H. Socio-political terms	12	—
I. Discourse particles and formulae (greetings etc.)	4	1
J. Exclamations and interjections	6	3

Table 1: Fijian and Hindi loans in Fiji English

Geraghty 1977; Thomson 1999), which regularly appear in the daily tabloids without a gloss, or are used in radio and television English language news broadcasts.

For each word respondents were required to indicate whether they: (a) definitely knew the meaning of the word but actually never used it, and (b) had used (or still used) the word in everyday English conversations with Fijians, Indo-Fijians or other expatriates. Words which respondents marked under category (a) were seen as part of the respondents' passive vocabulary, whilst words marked under (b) were seen as constituting part of their active vocabulary.

2.2 Participants

Responses were obtained from 140 expatriates (73 males, 67 females) from 20 different countries, three quarters of whom (75.6%) came from English speaking countries.[6] Respondents were found on a networking basis (Holmes, Bell & Boyce 1991: 23-25), and ranged in age from 12 to 68, with a mean age of 33 (see Table 2). Length of residence in Fiji ranged from six months to 23 years, with a mean length of residency of four years, nine months (see Table 3).

Education levels of respondents were as follows: secondary educated 35% (49/140), tertiary educated 65% (91/140). One third of respondents (33.6%) were engaged in full-time study either at secondary school (23.6%) or university (10%), the rest were either engaged in domestic duties (12.9%) or in some professional occupation (53.5%).

AGE GROUP	NUMBER & PERCENTAGE OF SAMPLE
10 - 20 years	34 (24.3%)
21 - 30 years	21 (15.0%)
31 - 40 years	39 (27.9%)
41 - 50 years	23 (16.4%)
51 - 60 years	20 (14.3%)
61 - 70 years	3 (2.1%)

Table 2: Age distribution of sample

LENGTH OF RESIDENCE	NUMBER & PERCENTAGE OF SAMPLE
< 1 year	15 (10.7%)
1 - 3 years	60 (42.9%)
4 - 6 years	36 (25.7%)
7 - 10 years	12 (8.6%)
> 10 years	17 (12.1%)

Table 3: Distribution of length of residence

3. Results

Table 4 shows the mean number of Fijian and Hindi words known and used with their standard deviations. The large standard deviations reflect the heterogenous population sample, as it consisted of a wide variety of ages, length of residence, occupations, and countries of origin.

Although Bellett's data cannot realistically be compared with this data, it is nonetheless, at least worth comparing both studies' results for the mean number of words known and used. The mean number of Māori words Bellett's respondents knew and used were 41 and 26 respectively. This shows her respondents had a larger passive than active Māori vocabulary. The reverse is true for respondents in my survey, where Fijian and Hindi active vocabularies were larger than passive vocabularies. This is likely to be due to the pragmatically different linguistic environments and circumstances each group of respondents find themselves in. Given the limited time expatriates normally stay in Fiji, not to mention the general lack of incentive to learn more than a minimum of Fijian or Hindi words, it is unlikely they will learn many of them

LOANWORDS	MEANS	STANDARD DEVIATIONS
Fijian words known (out of a total of 98)	13.5	10.8
Fijian words used	29.4	15.4
Fijian words known & used combined	41.6	18.4
Hindi words known (out of a total of 45)	3.6	3.3
Hindi words used	5.3	3.3
Hindi words known & used combined	7.6	5.9
Fijian & Hindi words known combined (out of a total of 143)	16.1	12.8
Fijian & Hindi words used combined	34.2	18.8
Fijian & Hindi words known & used combined	49.7	22.3

Table 4: Mean number of words known and use

unless they are pragmatically useful or are required in every-day conversations.[7] On the other hand, New Zealand Pākehā now seem to be under growing pressure to learn, or at least be familiar with a wide range of Māori terms. This is a consequence of New Zealand's attempts to move closer to being a bicultural society. In this context, the Māori language and culture have enjoyed an increase in both status and esteem over the last decade or two. The increase in educational initiatives to promote both Pākehā and Māori knowledge of the Māori language may also be a corollary of this.

Table 5 shows the combined percentage rates (in descending order) for declared knowledge of meanings for Fijian words and their use.[8] The category of loan (as enumerated in Table 1) is also indicated for each word .

The mean number of Fijian words used or known was 41.6 (42.4% of the 98 Fijian words). The vast majority of Fijian words that have less than a 30% combined use/knowledge rate have a higher rate of knowledge than use and may, therefore, be considered as part of respondents' passive vocabulary. Almost all words that have *more* than a 30% combined use/knowledge rate have a higher rate of use than mere knowledge of the meaning of the word, and may, therefore, be considered as forming part of expatriates' active Fijian vocabulary. The number of Fijian words used and known for individual respondents ranges from four (4.1%) to 88 (89.8%) words.

With the exception of the exclamatory approbation *uro* (which is rarely seen in print), all Fijian words with more than a 90% combined use and

Table 5: Combined response rates in descending order for Fijian loans

LEXICAL ITEM & CATEGORY OF LOAN		KNOW/USE %
bula 'hello'	**I**	100
sulu 'sarong'	**F**	100
dalo 'taro'	**B**	99
vinaka 'thank you'	**I**	98
lovo 'earth oven'	**F**	98
yaqona 'kava'	**E**	97
moce 'good-bye; good night'	**I**	97
bure 'house, hut'	**F**	96
meke 'k.o. dance'	**G**	96
uro 'exclamatory approbation'	**J**	93
yadra 'good morning'	**I**	84
walu 'kingfish'	**D**	83
tabua 'whale's tooth'	**G**	81
kokoda 'marinated raw fish'	**E**	81
tanoa 'carved wooden kava bowl'	**F**	80
lali 'hollowed out log drum'	**F**	76
sevusevu 'gifts presented at arrival'	**G**	75
isa! 'exclamation of surprise, sadness'	**J**	73
lolo 'coconut cream'	**B**	72
io 'yes'	**I**	71
sega 'no'	**I**	70
turaga 'chief'	**H**	69
masi 'bark cloth'	**F**	68
bilo 'cup made from _ coconut shell'	**F**	68
vanua 'land; region; community'	**H**	67
kumala 'sweet potato'	**B**	67
bilibili 'bamboo raft'	**F**	64
sa! 'expression of surprise'	**J**	64
kai 'freshwater mussel'	**D**	62
mataqali 'a kin group'	**H**	61
sasa 'k.o. hand-held broom'	**F**	61
bele 'leafy green vegetable'	**B**	61

(Table 5 continued)

LEXICAL ITEM & CATEGORY OF LOAN		KNOW/USE %
duruka 'wild sugar cane'	**B**	60
tilou/tulou 'pardon me'	**I**	58
tui 'king; chief'	**H**	57
talanoa 'a chat, informal talk'	**F**	52
vakalolo 'k.o. pudding'	**F**	49
kaiviti 'an indigenous Fijian'	**F**	48
waka 'most prised part of kava root'	**B**	47
rourou 'taro leaves cooked in lolo'	**B**	46
koro 'village'	**F**	45
vudi 'plantain banana'	**B**	43
drua 'twin-hulled canoe'	**F**	41
magimagi 'sinnet'	**F**	40
loloma 'love; gift'	**G/I**	39
turaga-ni-koro 'village chief'	**H**	39
dakua 'k.o. tree; timber'	**A**	39
mana 'mud crab'	**F**	37
taukei 'land owner'	**H**	36
bu 'green coconut'	**B**	36
ota 'fern eaten as a vegetable'	**B**	35
tikina 'district'	**H**	35
barewa 'exclamatory approbation'	**J**	34
yaka 'k.o. tree; timber'	**A**	34
taralala 'k.o. dance'	**F**	33
vesi 'k.o. tree; timber'	**A**	32
kavika 'Malay apple'	**A**	32
kerekere 'to cadge'	**G**	31
miti 'coconut cream with lemon juice'	**E**	31
kutu 'louse'	**F**	31
Bose Levu Vakaturaga 'Great Council of Chiefs'	**H**	31
ivi 'Tahitian chestnut'	**B**	30
salusalu 'a garland'	**F**	30
sobo 'expression of disapproval'	**J**	28
kula 'collared lory; fringe around a mat'	**C**	28

(Table 5 continued)

LEXICAL ITEM & CATEGORY OF LOAN		KNOW/USE %
reguregu 'condolence gathering'	**G**	27
colo 'the bush, highlands'	**F**	27
wi 'Tahitian apple'	**B**	26
soli 'k.o. fundraising event'	**G**	24
lewena 'inferior part of kava root'	**B**	24
malua 'later, bye-and-bye'	**F**	23
yaya 'stuff, things, belongings'	**F**	22
kana 'to eat; food, a meal'	**F**	21
uvi 'yam'	**B**	21
vasu 'a part-European; nephew'	**H**	20
cibi 'pre-match war song'	**G**	19
saqa 'trevally'	**D**	19
kailoma 'a part-European'	**F**	19
mokusiga 'to hang about'	**F**	19
kanikani 'skin condition from too much kava'	**F**	18
yavusa 'largest kinship group'	**H**	16
kati 'k.o. card game'	**F**	16
rara 'village green; grassed area for meetings'	**F**	14
taura tale 'village hoe down'	**F**	14
leqa 'trouble'	**F**	14
uto 'breadfruit'	**B**	13
magiti 'a feast'	**F**	13
voivoi 'pandanus leaves'	**A**	12
qari 'mud crab'	**D**	12
vulagi 'a stranger'	**F**	11
Vola ni Kawa Bula 'Fijian Register'	**H**	11
gunu sede 'k.o. fundraising event'	**G**	11
balolo 'k.o. annelid'	**D**	10
kosa 'kava dregs'	**F**	10
teitei 'vegetable garden'	**F**	8
vakasoso 'k.o. desert'	**E**	8
kuro 'cooking pot'	**F**	4
gatu 'large printed bark cloth'	**F**	1

Table 6: Combined response rates in descending order for Hindi loans

LEXICAL ITEM & CATEGORY OF LOAN		KNOW/USE %
oti 'unleavened bread'	**E**	95
samosa 'deep-fried savoury pastry'	**E**	89
puri 'small deep-fried flat bread'	**E**	61
bhindi 'okra/ladies fingers'	**B**	51
bhaji 'k.o. green leafy vegetable'	**B**	44
acha 'okay, fine, good, great'	**I**	34
girmit(*iya*) 'indenture period'	**F**	31
achar 'relish made from green fruits'	**E**	31
barfi 'k.o. sweet'	**E**	27
bhuja ' salty snack of peas, peanuts etc'	**E**	25
lakri 'k.o. deep-fried sweet'	**E**	25
bara 'k.o. deep-fried savoury'	**E**	18
dhaniya 'coriander'	**B**	16
murgee/murga 'a chicken'	**D**	15
junglee 'wild; uncivilised person'	**F**	15
sirdar 'foreman, overseer'	**F**	15
gulgula 'k.o. deep-fried sweet'	**E**	15
baigan 'eggplant'	**B**	14
seo 'k.o. savoury'	**E**	14
mattar 'savoury peas'	**E**	12
choro 'to steel'	**F**	12
jalebi 'k.o. sweet'	**E**	12
kedgeree 'dish of rice, dhal & onions'	**E**	24
puja 'Hindu religious rite; prayer'	**G**	21
kisan 'farmer'	**F**	20
halwa 'k.o. sweet'	**E**	19
paisa 'money'	**F**	11
ghazal 'slow, emotional love song'	**G**	11
jira/jeera 'cumin seed'	**B**	9
sarso 'mustard seed'	**B**	9
bhajan 'Hindu devotional song'	**B**	8
hardi/haldi 'tumeric'	**B**	8

(Table 6 continued)

LEXICAL ITEM & CATEGORY OF LOAN		KNOW/USE %
piala 'small enamel bowl'	**F**	8
bhaia/bhaini 'brother; male friend'	**F**	8
brinjal 'eggplant'	**B**	6
chor 'a thief'	**F**	5
paidar 'to go on foot'	**F**	4
suji 'semolina'	**E**	3
trup 'k.o. card game'	**F**	3
chauraiya 'amaranth spinach'	**B**	2
tulsi 'sweet basil'	**B**	1
chamar 'a ne'er-do-well'	**F**	1
tawa 'iron plate for cooking rotis'	**F**	0
katha 'prayer ceremony'	**G**	0
pakora 'k.o. savoury snack'	**E**	0

knowledge rate are seldom glossed or italicised in the English language print media. It is also worth pointing out that three of the five Fijian words to have found their way into the *OED* (*sulu* 'sarong', *yaqona* 'kava', and *bure* 'house') are among the top ten words. These three words, along with *ivi* 'Tahitian chestnut' (which is also listed in the *OED*), are indicated in bold type. Another three of the top ten words are the discourse particles *bula*, *moce*, and *vinaka* 'thank you'. Most tourists who leave Fiji after a week's holiday will be quite familiar with all of these six words.

Table 6 shows the combined percentage rates (in descending order) for declared use and knowledge of meanings for Hindi words. The category of loan is also indicated for each word (see Table 1). Hindi items that are also listed in the *OED* are shown in bold type. The mean number of Hindi words used and known is 7.6 (15.5% of the 45 Hindi words), and ranged from zero to 28 (62.2%) words.

A similar result for 'active' vs 'passive' vocabulary seen for Fijian words was obtained for Hindi words. All Hindi words that have *more* than a 30% combined use and knowledge rate have significantly higher rates of use than mere knowledge of their meanings. These words may be considered forming part of expatriates' active vocabulary. On the other hand, the majority of words

with *less* than a 30% combined use and knowledge rate have significantly higher rates of knowledge of meaning than rates of use. According to the paradigm, these words should be seen as constituting part of expatriates' passive Hindi vocabulary.

When both the Fijian and Hindi loans are considered together, the average number of words used or whose meaning is known is 49.7 (34.3% of all Fijian and Hindi words).

Since the large standard deviations (s.d.) for the mean number of words known and used is indicative of the heterogenous nature of the population sample, it is worth examining responses in terms of gender, age, and length of residence.

3.1 Gender

Generally, there is no statistically significant difference between males and females in their use of and familiarity with the meanings of Fijian and Hindi words, (t = -.42, df = 138, 2-tailed p = .673).[9] This result coincides with Bellett (1995:82) who found no significant gender difference in knowledge and use of Māori words amongst her New Zealand English speaking respondents. The mean number of Fijian and Hindi words known by males is 48 (s.d. = 25), and for females 50 (s.d. = 20).

For some individual words, however, there are statistically significant differences between males and females. Females tend to use or know the meaning of most Fijian and Hindi names for fruit, vegetables, prepared dishes, and household items much more than males (e.g. *uvi*, *bele*, *lolo*, *ivi*, *vakalolo*, *wi*, *kokoda*, *kavika*, *bu*, *duruka*, *kumala*, *sasa*, *roti*, *bhaji*, *baigan*, *halwa*, *sarso*, *lakri*, *jira*, *barfi*, *gulgula*). The men, on the other hand, clearly showed more knowledge of Fijian socio-political terms (e.g. *vanua*, *reguregu*, *Bose Levu Vakaturaga*, *matagali*, *taukei*, *tikina*) kava terminology (e.g. *waka*, *lewena*, *kosa*), the names of different types of timber (e.g. *dakua*, *vesi*), colloquialisms, exclamations and interjections (e.g. *barewa*, *uro*, *acha*). These specific gender differences are perhaps indicative of the different social roles fulfilled by many expatriate men and women in Fiji. If females who are engaged in full-time study (i.e. school students) are disregarded (21/67 = 31.3%), there are 28 (41.8%) who were employed in professional occupations and 18 (18.9%) who were engaged in domestic duties and the buying and cooking of the family's meals. Even women who were in full-time employment would have some, if not most, influence on the buying and preparation of the family's meals. Women's greater knowledge and use of Fijian and Hindi

names for foods and dishes is, therefore, not all that surprising. The 50 males (50/73 = 68.5%) who were engaged in professional employment outside the home would tend to have a broader range of business and business related social contacts, than most females in the sample. In the Fiji business world, the use of socio-political terms is quite commonplace.

3.2 Age

Age plays a significant factor in the use and knowledge both Fijian and Hindi words. A 1-Way Anova revealed that the older the respondent, the more likely he/she knew the meanings of Fijian and Hindi words and used them in day-to-day English interactions. Table 7 shows the mean number of Fijian and Hindi words (combined) used and known for each age group.

This trend is reversed with use and knowledge of colloquialisms and exclamations (e.g. *sa*, *sobo*, *barewa*, *uro*, *mokusiga*, *yaya*, *paidar*, *choro*, *paisa*), and the names of Indian sweets (e.g. *barfi*, *gulgula*, *jalebi*, *lakri*). In all these cases, teenagers (the majority of whom were secondary students at Suva's International School) know and use them much more than any other age group. Once again, this is not that surprising as expatriate children have access to and are exposed to colloquial and basilectal speech much more than people their parents' age. And above all, most young teenagers at the International School are known to relish Indian sweets which are sold by street vendors outside the school during lunch breaks.

Table 8 shows the mean number of Fijian and Hindi words used and known by each age group. It shows that the older the respondent, the more Fijian and Hindi words are known and used. The only exception is for Hindi

AGE GROUP	MEAN & (S.D.)
10 - 20 years	41 (18)
21 - 30 years	44 (25)
31 - 40 years	47 (22)
41 - 50 years	52 (26)
51 - 60 years	60 (16)
61 - 70 years	65 (9)

(F = 2.41, df = 139, p = .039)

Table 7: Mean number of Fijian and Hindi words (combined) used and known by age

AGE GROUP	HINDI WORDS MEAN & (S.D.)	FIJIAN WORDS MEAN & (S.D.)
10 - 20 years	8 (7)	36 (19)
21 - 30 years	5 (5)	37 (14)
31 - 40 years	6 (5)	40 (18)
41 - 50 years	7 (5)	46 (22)
51 - 60 years	7 (6)	50 (13)
61 - 70 years	11 (4)	58 (4)
	(1-Way Anova: F = 2.64, df = 5 & 135, p = .026)	(1-Way Anova: F = 2.88, df = 5 & 134, p = .017)

Table 8: Mean number of Fijian and Hindi words used and known by age

words in the 10-20 age group. This anomaly is due to that age group's familiarity with Indian sweets.

The overall steady increase in use and knowledge of Fijian and Hindi loanwords as the age of the respondent increases is intriguing. Research findings are inconclusive in regard to the influence of age related factors on the acquisition of a second language (Hatch 1983:188-197). Weighed collectively, published research findings cannot strongly support an optimal age hypothesis that says 'the younger the better', nor can they embrace the contrary hypothesis 'the older the better'. Since my simple survey was not designed to examine age related factors in the acquisition of a foreign lexis, no appropriate data was collected during the survey that could shed light on the matter. It would, therefore, be unwise to speculate upon the reasons for this finding.

The possibility that age and length of residence may be correlated in terms of use and knowledge of Fijian and Hindi words was also examined. Table 9 shows that the mean age for respondents who have lived in Fiji for more than one year does not differ significantly (1-way Anova: F = 1.215, df = 3 & 121, p = .31).

3.3 Length of Residence

As might be expected, the number of years an expatriate lives in Fiji has a profound effect upon the use and knowledge of Fijian and Hindi words. The longer an expatriate lives in Fiji, the more Fijian and Hindi words are known

PERIOD OF RESIDENCE	MEAN AGE & (S.D.)
< 1 year	25 (12)
1 - 3 years	31 (12)
4 - 6 years	37 (13)
7 - 10 years	34 (19)
> 10 years	37 (21)

(F = 2.39, df = 4 & 135, p = .05)

Table 9: Mean age for each period of residence

and used. This trend is clearly shown by the mean number of Fijian and Hindi words used and known for each period of residence (Table 10).

The steep learning curve for Fijian and Hindi words over the first seven years steadily levels off as length of residence increases. After seven years' residence few words seem to be added to the expatriates' lexicon. The effect is analogous to the figures of five to seven years required for attainment of the full range of second language acquisition among a heterogenous L1 population in Canada (Cummins 1984), as well as the levelling off in English proficiency among Mexican immigrants after approximately eight years' residence in California (Hakuta & D'Andrea 1992).

When Fijian and Hindi words are considered separately (Table 11), the same trends are seen as in Table 10. There is a steady increase in the number of words the expatriate acquires for the first seven years of residence, followed by a distinct levelling off of the number of words learnt.

PERIOD OF RESIDENCE	MEAN & (S.D.)
< 1 year	27 (13)
1 - 3 years	42 (21)
4 - 6 years	55 (21)
7 - 10 years	66 (21)
> 10 years	67 (12)

(1-Way Anova: F = 13.38, df = 4 & 135, p = .00001)

Table 10: Mean number of Fijian and Hindi words (combined) used and known by period of residence

PERIOD OF RESIDENCE	NO. HINDI WORDS MEAN & (S.D.)	NO. FIJIAN WORDS MEAN & (S.D.)
< 1 year	3 (2)	25 (12)
1 - 3 years	6 (5)	36 (17)
4 - 6 years	8 (6)	47 (17)
7 - 10 years	11 (6)	55 (16)
> 10 years	11 (5)	56 (11)
	(1-Way Anova: F = 7.803, df = 4 & 135, p = .00001)	(1-Way Anova: F = 12.51, df = 4 & 134, p = .00001)

Table 8: Mean number of Fijian and Hindi words used and known by age

4. Concluding remarks

The main outcomes of this small survey, which are ultimately not all that surprising, are:

- the Fijian and Hindi words which form part of expatriates' active vocabularies are also those that occur most often in the media and in conversation. These are the most widely known and recognised lexical items in Fiji English. Many of these words (especially discourse particles, names of edible flora and fauna, and those referring to important aspects of Fijian and Indian culture) are also pragmatically essential for effective day-to-day communication in Fiji. [10]
- The differences in the types of words used and known by women, men and teenagers reflect the immediate pragmatic value they have for each of these groups.
- The number of Fijian and Hindi words an expatriate acquires increases over the first seven years of residence, after which there is a distinct levelling off of the learning rate.

The process of lexical nativisation is a continuum, and therefore, the point at which an item may or may not be considered fully integrated into the recipient language is debatable and naturally arbitrary. As I mentioned above, this is not an issue for the speakers of Fiji English who are often quite conversant in each other's first languages. Applying expatriates' knowledge

and use of Fijian and Hindi words to gauge the degree of nativisation of these words is therefore not altogether appropriate. Fiji English does not 'belong' to English-speaking expatriates, but to the Fiji Islanders themselves. Which Fijian and Hindi words Fiji Islanders use (and how often) in their daily English interactions, is ultimately a matter for themselves to decide. Nevertheless, the use of these words by expatriates in *their* daily English interactions offer some insight into which Fijian and Hindi words have become fully nativised into Fiji English.

Notes

1 Fiji Hindi is a local variety of Hindi, which has evolved from a koine of various dialects of the Hindustani lingua franca of North India (Siegel 1987: 187-203).

2 Words borrowed from language x into language *y,* where they became nativised and underwent semantic shift, after which they were re-introduced into language *x*.

3 There are also quite a number of other first languages which include Rotuman, Cantonese, and Gilbertese.

4 That is rotis on the main front lawn of the university.

5 Fijian words are given in the usual Fijian orthography:

Symbol	IPA	Symbol	IPA
b	^{m}b	j	tʃ
d	^{n}d	c	ð
q	^{ŋ}g	v	β
g	ŋ		

6 North America (22.9%), Australia (18.6%), New Zealand (17.9%), United Kingdom (16.2%), Other Pacific islands (10.3%), Asia (7.8%), Africa (3.5%), Europe (2.8%).

7 This must also be seen in the light of the vastly differing population distributions of both countries. New Zealand is 70% Pākehā (white, European), whilst Fiji's expatriate population is 0.76% (0.4% European, 0.36% others).

8 The two variables, words that are known and words that are used, were combined for ease and clarity of analysis. Moreover, the two variables are dependent upon each other. Since use is dependent upon knowledge (use is ultimately restricted by knowledge), items in the active vocabulary form a subset of the passive vocabulary.

9 For Fijian words (t = -.17, df = 136, 2-tailed p = .86), for Hindi words (t = -1.08, df = 138, 2-tailed p = .28).

10 On 6 November 1999, the *Fiji Times* published the first of its weekly Fiji Words crossword puzzles which include clues on Fiji culture, history and geography. The first seven crosswords included the following words in either the clues or

solutions: *qio* 'shark', *masi*, *sulu*, *tanoa*, *tulou*, *vinaka*, *teitei*, *tui*, *moce*, *baigani*, *sirdar* and *roti*. The ratio of Fijian to Hindi words used in the crosswords reflects that of the Fijian and Hindi loans in the lexis of Fiji English.

References

Arms, Fr. David G. 1975. 'The Church and the Vernacular'. *Outpost: A Fiji Forum* 3 (6): 2-19.

Bellett, D. 1995. 'Hakas, hangis, and kiwis: Måori lexical influence on New Zealand English'. *Te Reo* 38: 73-104.

Cummins, J. 1984. *Bilingualism and Special Education: Issues in Assessment and Pedagogy*. San Diego: College-Hill Press.

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

Geraghty, Paul. 1977. 'Fiji pidgin and bilingual education'. *Fiji English Teachers' Journal* 12: 2-8.

Hakuta, K. & D. D'Andrea. 1992. 'Some properties of bilingual maintenance and loss in Mexican background high-school students'. *Applied Linguistics* 13 (1): 72-99.

Hatch, E.M. 1983. *Psycholinguistics: A Second Language Perspective*. Rowley, Mass: Newbury House.

Holmes, Janet, Allan Bell and Mary Boyce. 1991. *Variation and Change in New Zealand English: A Social Dialect Investigation*. Wellington: Department of Linguistics, Victoria University, Wellington.

Kelly, Sr. Francis. 1975. 'The English Spoken Colloquially by a Group of Adolescents in Suva'. *Fiji English Teachers' Journal* 11: 19-43.

Moag, Rodney F. and Louisa B. Moag. 1977. 'English in Fiji, Some Perspective and the Need for Language Planning'. *Fiji English Teachers' Journal* 13: 2-26.

Monsell-Davis, Michael. 1984. 'Current slang expressions among young people in Suva, Fiji'. Unpublished typescript.

Siegel, Jeff. 1987. *Language Contact in a Plantation Environment: A Sociolinguistic History of Fiji*. Cambridge: Cambridge University Press.

Siegel, Jeff. 1989. 'English in Fiji'. *World Englishes* 8 (1): 47-58.

Thomson, Peter. 1999. *Kava in the Blood*. Auckland: Tandem Press.

A DIACHRONIC ANALYSIS OF NON-STANDARD VOCABULARY IN THE DIALOGUES OF 1930S FRENCH FILMS

Michaël Abecassis: *Oxford University Language Centre, Oxford University (46 Southmoor Road, Jericho, Oxford, OX2 6RD)* <*michael.abecassis@modern-languages.oxford.ac.uk*>

Abstract

This study examines non-standard lexical items in a 1930s corpus of French films. Its purpose is to investigate the evolution of such items and gauge whether they are still part of today's linguistic repertoire of French people. The study is based on an analysis of French dictionaries prior to and contemporary with the films which have been compared with today's to assess the degree of acceptability of the words investigated, as well as on surveys conducted in 1999 on a sample of French speakers.

1. Introduction

The existence of early sound recordings enables us to study aspects of language change in French over the 20th century in a way that was not possible in earlier centuries. In this analysis[1] I investigate the evolution of lexicographers' attitudes towards non-standard items found in a corpus of 1930s films, over a sixty-year period. For this purpose, I compiled a glossary of the non-standard words in the films used. In the first section, I ask whether the colloquial vocabulary used in the films investigated is still in use today, by comparing the labels of words in *Le Larousse du XX*e *siècle* (1932) with first of all those

used by *Le Nouveau Petit Robert* (1993) and secondly by *Le Petit Larousse* (1989). I would like to mention at this stage that the former is less conservative in its policy than the latter, which certainly has an effect on the results. In the second part, I estimate the vitality of non-standard vocabulary through the perception of current users of the French language.

2. The Corpus

For this research, a corpus of French films (recorded on videocassette) dating from the 1930s has been assembled. This provides interesting and previously unexploited evidence concerning Parisian vernacular speech at that time. The film corpus comprises five black and white films: *Hôtel du Nord* (1938), *Fric-frac* (1939), *Circonstances atténuantes* (1939), *Le Jour se lève* (1939), *La Règle du jeu* (1939).

Adapted from a play by Bourdet, *Fric-frac* stars Arletty, Michel Simon and Fernandel. The film derives much of its humour from the contrast between the 'vernacular speech' of the proletarian speakers and the 'standard speech' of the upper-group characters. Fernandel, acting as a bridge between the two, attempts to integrate himself with petty thieves by accommodating to them both socially and linguistically. The comic film *Circonstances atténuantes*, starring Michel Simon, Suzanne Dantès, Dorville and Arletty, presents a variety of Parisian speakers from different social classes. *Le Jour se lève* encapsulates the pessimistic mood that pervaded France in the 1930s. In Carné's film, scripted by Jacques Prévert, Jean Gabin portrays the tormented working-class male, who murders the obnoxious seducer of an innocent flower-seller, played by Jules Berry. *La Règle du jeu* is one of Jean Renoir's most famous films. It is a satire and a caricature of the bourgeoisie in its most stereotyped form. Marcel Dalio and Mila Parély play the aristocrats, while Julien Carette is the working-class Parisian speaker. Finally, Arletty's line 'atmosphère, atmosphère' contributed to the popularity of Carné's *Hôtel du Nord*. Jeanson, the author of this famous dialogue between Jouvet and Arletty, wrote the scripts of several classics which opened the way to poetic realism, offering 'les gens du peuple' important roles in 1930s French cinema.

I chose these films first because of their lasting popularity. They are some of the most famous films of the 1930s. Above all, however, I thought they were plausibly representative of the most stereotypical Parisian sociolect of that period.

3. Diachronic Analysis

3.1 French dictionary labels

The cleavage between the prestige norm (codified usage) and non-standard items (colloquial usage) in the lexicon can be expressed in terms of High and Low varieties. All languages seem to have a stock of high-value words reserved for formal contexts and low-value words for informal situations. In French dictionaries, the convention is to adopt such labels as *arg*. (*argot*, 'slang'), *enf*. (*enfantin*, 'childlike'), *fam*. (*familier*, 'informal'), *pop*. (*populaire*, 'lower class'), *rég*. (*régional*, 'regional'), *triv*. (*trivial*, 'crude'), *vieilli* ('outdated'), *vulg*. (*vulgaire*, 'vulgar') and *vx* (*vieux*, 'obsolete') etc. *Familier* is often confused with *populaire*. The former is, as Battye and Hintze note, 'a register' that is mostly associated with informal situations (1992: 340) and spontaneous French. Constructions, expressions and lexis that are stylistically marked as *familier* allegedly do not belong to a definite socio-economic group but are shared by both lower and upper-class speakers. The label *populaire* is prescriptively thought to convey a negative sociological connotation in contrast to the *français familier*. It concerns varieties of French that are primarily spoken.

3.2. Stylistic evolution of ten non-standard items

By way of introduction, the following study examines the treatment of ten examples of non-standard lexical items found in my corpus by the best-established dictionaries from the 17th to the 20th centuries, by looking at their stylistic labels. A similar exercise was carried out by Désirat and Hordé in 1976 and later by Mūller (1985), but over a much shorter period of time, focusing solely on 20th century dictionaries. Désirat and Hordé's findings of language levels in three different dictionaries illustrated that '*Le Petit Larousse* (1970) [était] plus conservateur que *Le Littré* et le *Petit Robert* plus libéral que ses prédécesseurs' (1976: 43-44).[2]

The purpose of this exercise is firstly to establish when these words were first lemmatised and to see how their level of acceptability has changed throughout the centuries, and secondly, to see whether there is much variation between dictionaries. The following words were chosen at random from the glossary.

Table 1 does not feature any work prior to 1932, as none of the prominent French dictionaries examined, ranging from Furetière's *Dictionnaire universel* (1690) and the first edition of the *Dictionnaire de l'Académie* (1694) to the

Dictionaries	pépère	pèze	picolo	piger	pinard	pognon	poireau	popotin	potasser	proprio
Bloch O. and Wartburg W. von, 1932, *Dictionnaire étymologique de la langue française*, Paris, PUF	-	-	-	terme familier	argot scolaire	-	-	-	terme familier	-
Dictionnaire de l'Académie, 1931-5, 2 vol., 8th edition, Paris, Hachette	-	-	-	-	pop.	-	-	-	abs.	-
Le Larousse du XX^e siècle, 1932, Paris, Larousse	enf.	fam.	pop.	-	arg. militaire	arg.	-	pop.	arg. des écoles	arg. parisien
Dauzat A., Dubois J. and Mitterand H., 1964, *Nouveau dictionnaire étymologique et historique*, Paris, Larousse	fam.	pop.	pop.	pop.	pop.	pop.	-	-	arg. scolaire	-
Dictionnaire du français contemporain, 1966, Paris, Larousse	pop.	-	-	pop.	pop.	pop.	fam.	-	fam.	-
Robert P., 1985, 9 vol., 2nd edition, *Le Grand Robert de la langue française*, Paris, Robert	pop.	fam.	pop.	fam.	pop.	fam.	fam.	pop.	fam.	-
Trésor de la langue française, 1988, 16^e volumes, Paris, Gallimard	pop.	arg. and pop.	vieilli, pop.	pop. and fam.	pop.	arg. and pop.	fam.	pop., fam.	fam.	pop.
Dictionnaire Larousse, 1989, Paris, Larousse	fam.	arg.	-	fam.	pop.	pop.	fam.	très fam.	fam.	pop.
Le Nouveau Petit Robert, 1993, Paris, Robert	fam.	arg.	fam., vieilli	fam.	fam.	fam.	-	fam.	fam.	fam.

Table 1

eighth edition (1931-35) recorded these words. It has to be said that in early dictionaries, there were no labels in abbreviated form, and style was hardly referred to at all. The inclusion of dictionaries in table 1 has therefore been considered in terms of the registers they embrace. Indeed, they have all conventionally adopted the same labels to indicate degrees of informality: *fam*., *pop*., *arg*. and *vulg*. Table 1 shows that the non-standard items investigated only started to appear in general dictionaries of French in the 1930s. However, this does not necessarily mean that they could not be found in earlier literature. The words *piger* and *pognon*, for example, are mentioned in Sainéan's glossary to *Les Sources de l'argot ancien* (1912). It is the frontier between *fam*. and *pop*. that dictionaries find most difficult to define. One can note that the label *pop*. confuses stylistic and social criteria, whereas *fam*. is purely stylistic. Over the past sixty years, the tendency is for *pop*. words to become *fam*. and for *arg*. words to become *pop*. The attribution of stylistic labels in *Le Petit Larousse illustré* (1989) differs from that of *Le Robert* (1993) as regards the labels *pop*. and *fam*., but there is agreement concerning the label *arg*.

Before correlating *Le Larousse du XX*^{e} *siècle* (1932) with *Le Petit Larousse Illustré* (1989), it was felt necessary to check the representativity of *Le Petit Larousse Illustré* by comparing its treatment of the non-standard words in the film corpus with that of *Le Nouveau Petit Robert* (1993). The latter is generally regarded as not excessively prescriptive (Lodge 1989: 430).

Le Petit Larousse appears to be more conservative than *Le Petit Robert*. The disagreement between the two dictionaries is found in the percentage of

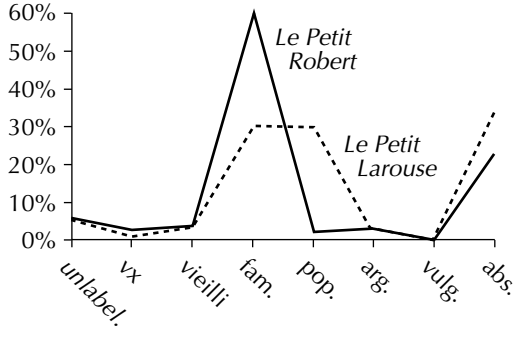


Figure 1

the words that receive the labels *fam*. or *pop*. Most of the lexical items allocated the style-label *pop*. by *Le Petit Larousse* appear with the label *fam*. in *Le Petit Robert*. *Le Petit Robert* obtains a lower score of *absent* words[3], which suggests a greater readiness to accept these items into the general language. There seems to be agreement between lexicographers on the words labelled *arg*.

3.3 Comparison of style-labels in Le Larousse du XXe siècle (1932) with Le Petit Larousse Illustré (1989)

In what follows I will consider the changes which have taken place between 1932 and 1989 in the way lexicographers regard these words. I intend to compare like with like by studying two different editions of Larousse dictionaries. A preliminary caveat concerns the relative sizes of the dictionaries. The *Larousse du XX*e *siècle*, in six volumes, compiles 120,000 words (Matoré 1968: 142), while *Le Petit Larousse illustré* has 58,000. Despite the difference in the number of words, the two dictionaries work in the same style. Paul Augé in his introduction to *Le Larousse du XX*e *siècle* (1932: 1) claims that the dictionary is descriptive:

> Ainsi, à l'ancien fonds de la langue, dont les Dictionnaires Larousse ont toujours fourni l'inventaire le plus exact, nous avons ajouté tous les mots nouveaux, sans négliger les termes spéciaux ou techniques, ni même les mots d'argots, réalisant de la sorte dans nos colonnes le vrai miroir du "français vivant"'.[4]

However, in spite of this statement, *Le Larousse du XX*e *siècle* is highly prescriptive as this study shows.

In these dictionaries, some words appear *unlabelled* (abbreviated below as *unlabel*.), which I take to mean that the lexicographers consider these words to be *standard*. However, one dictionary might regard a word as *non-standard* and another not. The adjective *mimi,* 'cutie' standing for *mignon* was *unlabelled* in 1932 and rated *fam*. in 1989. This shows that lexicographers' views about *standard* words are not necessarily stable.

Tables 2-3 and Figures 2-3 show the numbers of words labelled *vx, vieilli, enf., fam., pop. arg., triv*. and *vulg.,* in both *Le Larousse du XX*e *siècle* and in the *Le Petit Larousse illustré*, and those items that were excluded. Figure 4 indicates the evolution of the stylistic labels that the two dictionaries have in common. This exercise takes into account *types* (*members*) only and is restricted to the 450 items compiled in my glossary.[5] Labels are ranked in descending order of socio-stylistic value from the left to the right of the table.

LE LAROUSSE DU XXE SIÈCLE									
unlabel.	*vx*	*vieilli*	*enf.*	*fam.*	*pop.*	*arg.*	*triv.*	*vulg.*	*abs.*
3%	0%	0%	0.5%	11%	14%	17%	1%	0%	53.5%

Table 2

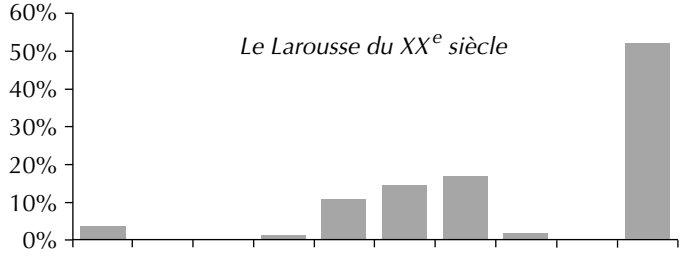


Figure 2

LE PETIT LAROUSSE ILLUSTRÉ								
unlabel.	*vx*	*vieilli*	*rég.*	*fam.*	*pop.*	*arg.*	*vulg.*	*abs.*
5%	1%	3%	0.1%	28.5%	28.3%	2.4%	0.1%	31.6%

Table 3

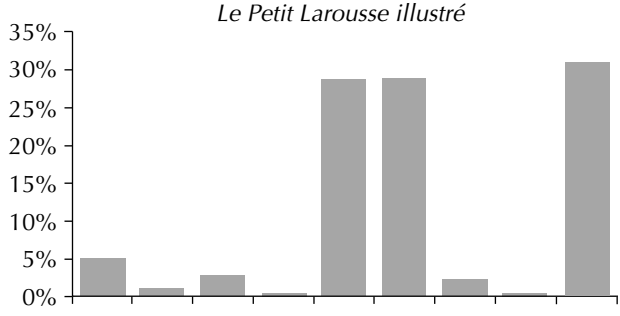


Figure 3

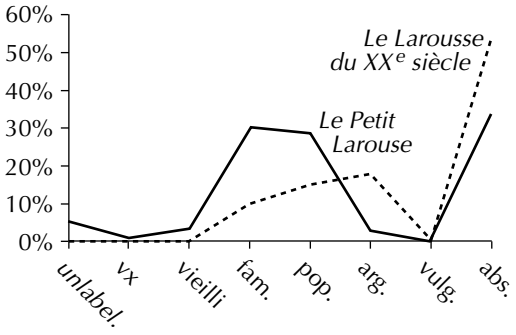


Figure 4: Style-labels common to *Le Larousse du XX*[e] *siècle* and *Le Petit Larousse illustré*

3.4 Conclusions and observations

Figure 4 gives us indications of the way attitudes to language have evolved over a sixty-year period. The most outstanding feature is that *Le Larousse du XX*[e] *siècle* excludes the majority of the investigated items which appear in my data. These lexical items were highly stigmatised in 1932 and attitudes to these items have evolved. They are now conventionally labelled *fam*., *pop*., or *arg*.. Figure 3 shows that a small ratio of *unlabelled* words in 1932 (*un gars*, 'a guy'; *faire marcher,* 'to pull somebody's leg'; *un métèque*, 'a wog'; *merde*, 'shit' etc.) have been placed in the *fam*. category today. On the other hand, a small percentage of words rated non-standard in *Le Larousse du XX*[e] *siècle* (5%) have now been accepted into the standard (e.g. *aguicheuse*, 'a prick-teaser'; *dragées*, 'bullets'; *un crème*, 'a small white coffee'). The proportion of *arg*. words has decreased, falling from 17% in 1932 to 2.4% in the more recent dictionary. This indicates that traditional *argot* terms have now passed into the general French vocabulary. Certain patterns are more or less systematic. It is noticeable, for instance, that most words *absent* from the 1932 dictionary receive the label *pop*. in 1989, which means that sixty years later they are still heavily stigmatised and attributed to low-status groups. One item *absent* from the 1932 dictionary belongs in 1989 to the *vulg*. category (*un salopard*, 'a bastard'). Those that were labelled *arg*. are generally labelled *pop*. in 1989. Items with the labels *fam*. and *pop*. in 1932 are quite stable, despite a weak tendency for *pop*. words to become *fam*. One can say that the non-standard vocabulary of my film corpus cannot be characterised as specific

to a particular time period and is still commonly used today. Indeed, only 4% of the non-standard words in my corpus are now regarded as obsolete (*vx*) or outdated (*vieilli*).

4.0. Surveys of the Use of the Non-standard Lexical Items in Today's French

4.1 Purpose

One can derive some idea of the vitality of these non-standard items today from dictionaries like *Le Larousse* and *Le Robert*, but there is always the possibility that lexicographers have a different view of the lexicon from 'ordinary speakers'. Following Lodge (1989) and Armstrong (1998), I conducted my own survey of speakers' reported actual use of non-standard words in my corpus leaving aside their sense of the words' stylistic value.

4.2 The survey

In this survey, I presented a list of a hundred items from the film corpus to native French speakers currently resident in Scotland or France. Some of the informants were unknown to the investigator and contacted by electronic mail. The use of electronic mail restricted the people surveyed to those who were computer literate, but this method opens new perspectives for self-reporting questionnaires on the current use of language. Informants were asked to '*Soulignez les mots qu'il vous arrive d'utiliser dans la conversation*'.[6]

Table 4 presents the hundred words or expressions from my film corpus that were investigated. To avoid the confusion of some of these words with possible homonyms, I gave, where necessary, their 'standard' labels (e.g. *battant* in the sense of 'coeur', *bavard* in the sense of 'avocat'). The stylistic indicators are those given by *Le Petit Larousse illustré* (1989):

The following results are derived from the labelling of *Le Petit Larousse illustré*:

-*unlabelled*: 7
-*vieilli* labelled words: 1
-*fam*. labelled words: 34
-*pop*. labelled words: 30
-*arg*. labelled words: 6
-*vulg*. labelled words: 1
-*abs*.: 21

ITEMS	LABELS IN PETIT
à la revoyure (idiom), *so long*	*pop.*
allumeuse (noun), *prick-teaser*	*fam.* and *péj.*
apéro (noun), *aperitif*	*pop.*
au béguin (idiom), *in love*	*fam.*
avoir le béguin pour (idiom), *to have a crush on*	*abs.*
avoir quelqu'un à la caille (idiom), *to have a grudge against someone*	*abs.*
avoir un petit pépin pour (idiom), *to have a crush on*	*abs.*
balles (noun), *francs*	*fam.*
battant (noun) [coeur], *ticker*	*abs.*
bavard (noun) [avocat], *mouthpiece*	*abs.*
biberonner (verb), *to drink*	*fam.*
bicot (noun), *wog*	*abs.*
bistrot (noun), *bar*	*fam.*
bobos (noun), *bruises*	*fam.*, or *langage enfantin*
boniche (noun), *skivvy*	*pop.* and *péj.*
ça boume (idiom), *everything goes well*	*pop.*
cambrousse (noun), *countryside*	*pop.* and *péj.*
cambriole (noun), *burglary*	*abs.*
carbure (noun), *dough*	*abs.*
cassement (noun), *burglary*	*abs.*
c'est du nanan (idiom), yum-yum	*vieilli*
c'est pas bézef (idiom), *there is not much of it*	*pop.*
c'est le bouquet (idiom), *that's the last straw*	*fam.*
c'est ta noce (idiom), *it's your lucky day*	*abs.*
charognard (noun), *skunk*	*fam.*
cochon (noun), *pig*	*fam.*
colback (noun), *neck*	*pop.*
comaque (adjective), *like that*	*abs.*
combine (noun), *shady scheme*	*fam.*
comme dab (idiom), *as usual*	*abs.*
contredanse (noun), *police summons*	*fam.*
copain (noun), *pal*	*fam.*
copine (noun), *female pal*	*fam.*
coucou (exclamation), *hello*	*unlabelled*
crème (noun), *small white coffee*	*unlabelled*
Croquenot (noun), *beetle crushers*	*fam.*
dégueulasse (noun), *disgusting*	*pop.*
dragée (noun) [balles de pistolet], *bullets*	*unlabelled*
écluser un godet (idiom), *to sink a drink*	*pop.*
en avoir sa claque (idiom), *to be fed up*	*pop.*
en avoir plein le bide, *to be fed up*	*unlabelled*
entraver (verb), *to understand*	*arg.*
être à la page (idiom), *to be up-to-date*	*fam.*
être en boule (idiom), *to be angry*	*unlabelled*
être en rogne (idiom), *to be angry*	*fam.*
faire des paillardes (idiom), *to have sex*	*abs.*
faire du gringue à (idiom), *to chat up*	*fam.*
faire la gueule (idiom), *to pull a long face*	*pop.*
faire sisitte (idiom), *to sit down*	*abs.*
fauché (adjective), *broke*	*fam.*
flopée (noun), *a whole bunch*	*fam.*
flouse (noun), *dough*	*arg.*
fric-frac (noun), *break-in*	*pop.*
fripouille (noun), *swindler*	*fam.*
gauloise (noun), *cigarette*	*unlabelled*
godasse (noun), *boot*	*pop.*
gonflé (adjective), *cheeky*	*pop.*
guincher (verb), *to dance*	*pop.*
jetée (noun) [argent], *hundred francs*	*abs.*
la der des ders (idiom), *the last of all*	*pop.*

ITEMS	LABELS IN PETIT	ITEMS	LABELS IN PETIT
la ramener (idiom), *to grumble*	*pop.*	rigolo (adjective), *funny*	*pop.*
la trouver mauvaise (idiom), *not to find that all funny*	*fam.*	roteuse (noun), *cheap champagne*	*abs.*
machin (noun), *what's-his-name*	*fam.*	rupin (noun), *rich*	*pop.*
marc (noun), *coffee grounds*	*unlabelled*	salaud (noun), *bastard*	*pop.*
marrant (adjective), *funny*	*pop.*	salopard (noun), *bastard*	*vulg.*
micheton (noun), *prostitute's client*	*arg.*	sécher (verb) [boire], *to sink a drink*	*abs.*
ne pas être fichu de (idiom), *not capable of*	*fam.*	sécottine (noun), *pain in the neck*	*abs.*
ne pas les attacher avec des saucisses (idiom) *to be very mean*	*fam.*	en avoir marre (idiom), *to be fed up*	*arg.*
oseille (noun), *dough*	*pop.*	se grouiller (verb), *to hurry*	*pop.*
pèze (noun), *dough*	*arg.*	se magner le train (idiom), *to hurry*	*pop.*
piger (verb), *to understand*	*fam.*	s'en ficher (idiom), *not to care*	*fam.*
pinard (noun), *plonk*	*pop.*	en avoir plein le dos (idiom), *to be fed up*	*fam.*
pognon (noun), *dough*	*arg.*	s'en foutre (idiom), *not to give a damn*	*pop.*
pompe (noun), *beetle-crusher*	*pop.*	s'en jeter un (idiom), *to sink a drink*	*pop.*
popotin (noun), *bum*	*pop.*	s'en laver les mains (idiom), *not to give a damn*	*abs.*
potasser, *to swot*	*fam.*	se planquer (verb), *to hide*	*fam.*
pote (noun), *mate*	*pop.*	se rincer (verb), *to drink*	*abs.*
radin (adjective), *mean*	*fam.*	truc (noun), *thingummy*	*fam.*
raffut (noun), *noise*	*fam.*	veine (noun) [chance], *luck*	*fam.*
raide (adjective) [sans argent], *broke*	*abs.*	verni (adjective), *lucky*	*fam.*

Table 4: List of lexical items ranked in alphabetical order

4.3 Parameters of the experiment

In these statistics, I attempt to correlate speakers' perceptions of their use of non-standard items with age and gender parameters in order to discover which generation uses most of the low-status items in my 1930s corpus.

A) Sex

The sex of informants is believed by Milroy (1992: 163-179) to be a parameter of variation independent of style and social class. Male speakers are believed to use more non-standard items and swear words than females.

b) Age

The informants were evenly divided into six different age groups (see Lodge 1989):

10-15, 20-29, 30-39, 40-50, 51+, 65+

I did not find informants between 15 and 20 years old.

c) Origin of the speakers

The majority of the informants, roughly 75%, live in Paris, although some of them have moved temporarily to Scotland to study or teach. The other 25% live in other parts of France (Alsace, Jura, Rhône Alpes).

d) Social classes

In the spirit of a self-report survey, I invited the informants to locate themselves in one of the four categories:

1. Upper class
2. Middle class
3. Lower class
4. Without profession (pupils, students, retired people)

I am aware that the latter category is a 'hybrid' distinction, as it does not account for level of education, which is the important factor in this study. Retired people could belong to the upper, middle or lower class, being educated or uneducated. However, I felt that this extra category was needed, as none of the pupils, students or retired people questioned considered themselves to belong to any of the first three. Table 5 indicates how many informants fall in each category.

e) Problems with self-reporting surveys

The limitations of self-report questionnaires are obvious. The skewing of information in such surveys may be accounted for by the speakers' concern to preserve, in Brown and Levinson's terms, their 'faces' (1987). There is a risk that the informants might understate or overstate their usage of non-standard terms to protect their public images. The methodological issues of the size of the sample and the quantity of data are also serious issues (see Butler 1985; Milroy 1987; Asher's section on 'data collection' 1994, vol. 2: 815-816 and Stubbs 1983: 223-4). I make no strong claims about the representativeness of this survey, but it gives some idea of the current vitality of the colloquial vocabulary used in my films from the 1930s.

	NUMBER OF INFORMANTS
Upper class	17
Middle class	23
Lower class	1
Without profession (pupils, students, retired people)	19
Total	60

Table 5

4.4 The results

The general observation one can make is that most of these items are still used by today's French natives; although a small number (e.g. *c'est du nanan*, *c'est pas bézef*, *roteuse*) are only used by few informants.

Table 6 and figure 5 represent the average proportion of non-standard words which the different categories of speaker admit as belonging to their active vocabulary.

	10-15	20-29	30-39	40-50	51+	65+
Male speakers	25%	36%	59%	61%	48%	40%
Female speakers	24%	55%	46%	53%	38%	32%

Table 6

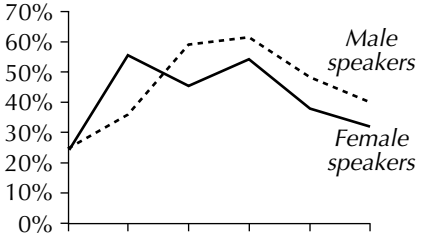


Figure 5

The percentages have been calculated according to the total number of lexical items listed (100). If an informant's perception of his use of the sampled non-standard items is 30 out of 100 then the percentage of items used will be 30%.

A) AGE GROUPS

The percentage of perceived use is smaller in the young than in the older speakers. Speakers between 10 and 15 years of age are the ones who use these non-standard idioms least. Most of the non-standard items investigated in this survey are not recognised by the younger informants. People in the 40 to 50-year-old category today, born in the 1950s, ten to twenty years after the films were first released, get the highest scores.

B) GENDER

In the 10-15 bracket, there is no clear difference between male and female speakers. Female speakers aged 20-29 obtain higher scores than males. Thereafter, the situation is reversed. This shows a tendency for females to use non-standard items as a means of expressing their identity as students (20-29) and to be more conservative in their speech after 30, when they start working.

C) SOCIAL CLASS

The scores for the different informants in their respective social classes are set out in table 7.

It would appear that there is no social correlation in the use of these items; the low score for the 'without profession' category is due to the number of informants between 10 and 15 years of age whose proportion of use is extremely low. It would be interesting to pursue this research further and to

	PERCENTAGE OF USE
Upper class	42%
Middle class	41%
Lower class	45%
Without profession (pupils, students, retired people)	33%

Table 7

assess whether 'ordinary speakers' have a different view from lexicographers, by questioning them on how they rate the non-standard words of the film corpus.

5. General conclusion

I carried out several diachronic studies, following the evolution of non-standard items of a 1930s film corpus over sixty years. The compilation of non-standard items into a glossary led to the comparisons of stylistic labels in a 1930s dictionary and a late 1980s dictionary, and permitted us to follow the evolution of stylistic indicators through time. General dictionaries have become increasingly tolerant of this kind of item. In pre-20th-century dictionaries, most of these words were absent. *Le Larousse du XX*e *siècle*, for instance, was highly prescriptive, and as a result most non-standard items were not included. Dictionaries from the 1980s tend to incorporate these non-standard words, but there is some disagreement on the stylistic label they should receive. On the whole, *Le Petit Larousse* appears to be less tolerant than *Le Petit Robert* and the frontier between *pop*. and *fam*. in dictionaries is particularly nebulous. The film script-writers clearly did not engage in any instances of lexical creativity, apart from very few one-off expressions. A survey among today's native French speakers revealed that a great deal of the vocabulary found in the film corpus was still commonly used in the 1990s and only rare expressions were rated as obsolete. It is also inferred from the statistics that it is middle-aged people who use the words of the film corpus most, but no social correlation can be established from these results. It is hoped that such an analysis will open further investigations on the language of the first talkies, which provide invaluable material for the understanding of the progresses of evolution of language.

Notes

1 The following revised research has been conducted in the course of my doctoral dissertation on "The Representation of Parisian Speech in the Cinema of the 1930s (University of St Andrews, Scotland 2000). I am very grateful to my supervisor Professor R.A. Lodge (University of St Andrews, Scotland) and to W. J. Anderson (University of St Andrews, Scotland) for her unstinting help when writing this paper.

2 '*Le Petit Larousse* (1970) [was] more conservative than *Le Littré* and *Le Petit Robert* more liberal than its predecessors'.
3 I use the label *abs*. standing for *absent* when a word is not found in the body of a dictionary.
4 'Thus, to the old reserve of language, the most accurate inventory of which has always been supplied by the *Larousse* dictionaries, I have added all the new words, without overlooking specialised and technical terms or even slang words, in this way achieving a true mirror of "living French" '.
5 *Tokens* (or *usages*) count the total number of words used by each character, therefore including multiple occurrences of the same word. Here, I disregard the frequency factor and count the total number of *word types* (or *members*) for each character.
6 'Underline words that you sometimes use in conversation'.

References

Armstrong, Nigel. 1998. 'La variation linguistique dans le lexique français'. *Zeitschrift für Romanische Philologie* 114: 462-495.
Asher, R.E. et al. (ed.) 1984. 9 vol. *The Encyclopaedia of Language and Linguistics*. Oxford: Pergamon Press Ltd.
Bloch, Oscar and Walther von Wartburg. 1932. *Dictionnaire étymologique de la langue française*. Paris: PUF.
Brown, Penelope and Levinson, Stephen C. 1987. *Politeness: Some Universals of Language Use*. Cambridge: Cambridge University Press.
Butler, Christopher Stuart. 1985. *Statistics in Linguistics*. Oxford: Blackwell.
Dauzat Albert, Dubois Jean and Mitterand Henri. 1964. *Nouveau dictionnaire étymologique et historique*. Paris. Larousse.
Désirat, Claude and Hordé, Tristan. 1976. *La Langue française au 20*e *siècle*. Paris: Bordas.
Dictionnaire de l'Académie. 1694. 2 vol. 1st edition. Paris: Coignard.
Dictionnaire de l'Académie. 1931-5. 2 vol. 8th edition. Paris: Hachette.
Dictionnaire du français contemporain. 1966. Paris. Larousse.
Furetière, Antoine. 1690. 3 vol. *Dictionnaire universel, contenant generalement tous les mots françois tant vieux que modernes*. La Haye et Rotterdam: A. and R. Leers.
Lodge, R. Anthony. 1989. 'Speaker's perception of non-standard vocabulary in French' in *Zeitschrift für Romanische Philologie*. 105 (5/6): 427-444.
Matoré, Georges. 1968. *Histoire des dictionnaires français*. Paris: Larousse.
Milroy, Lesley. 1987. *Language and Social Networks*. 2nd edition. Oxford: Basil Blackwell.
Milroy, Lesley. 1992. *Observing and Analysing natural language: a Critical Account of Sociolinguistic Method*. Oxford: Basil Blackwell.
Mūller, Bodo. 1985. *Le Français d'aujourd'hui*. Paris: Klincksieck.

Le Grand Robert de la langue française. 1985. 9 vol. 2nd edition. Paris: Robert.
Le Larousse du XX^{e} *siècle*. 1932. 6 vol. Paris: Larousse.
Le Nouveau Petit Robert. 1993. Paris: Robert.
Le Petit Larousse illustré. 1989. Paris: Larousse.
Le Trésor de la langue française. 1988. 17 vol. Paris: Gallimard.
Sainéan, Lazare. 1912. *Les Sources de l'argot ancien*. Paris: Champion.
Stubbs, Michael. 1983. *Discourse Analysis: the Sociolinguistic Analysis of Natural Language*. Oxford: Blackwell.

THE INFLUENCE OF THE MAORI POPULATION ON NZ DIALECT AREAS

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

Abstract

This paper examines data from a national survey of children's playground vocabulary, which has revealed evidence of significant dialect divisions in New Zealand. In particular, the Northern part of the North Island is distinct, and sometimes the North Island differs from the South. There is also evidence of considerable variation depending on socio-economic factors. This paper examines the hypothesis that the location of the Maori population in New Zealand is an important contributory factor in the patterns of regionalisation which have emerged.

1. Background

The results reported in this paper[1] derive from a study of the playground vocabulary of New Zealand school children. Year 7 and 8 students in 150 schools located from Kaitaia (in the far north of the North Island) to Bluff (at the Southern tip of the South Island) were surveyed by means of a questionnaire presented to them by their teacher. The distribution of the participating schools can be seen in the data maps presented below; 57 were in the South Island and 93 in the North Island.

The questionnaire covered the names of some playground games, playground rituals (e.g. what you say on the first of the month), basic social interactions, words for expressing feelings, and a few words for personal

stereotypes. Each teacher read out the scenarios in the questionnaire to a class of children and wrote down the children's responses. Multiple responses to questions were encouraged, and in most cases given, although a few teachers reported only majority forms. Our method treated all responses equally, and the shadings on the maps below indicate that the appropriate form was reported from that school. Thus the study was based on 150 sets of data, but there were sometimes as many as 20 responses to an individual question from one school.

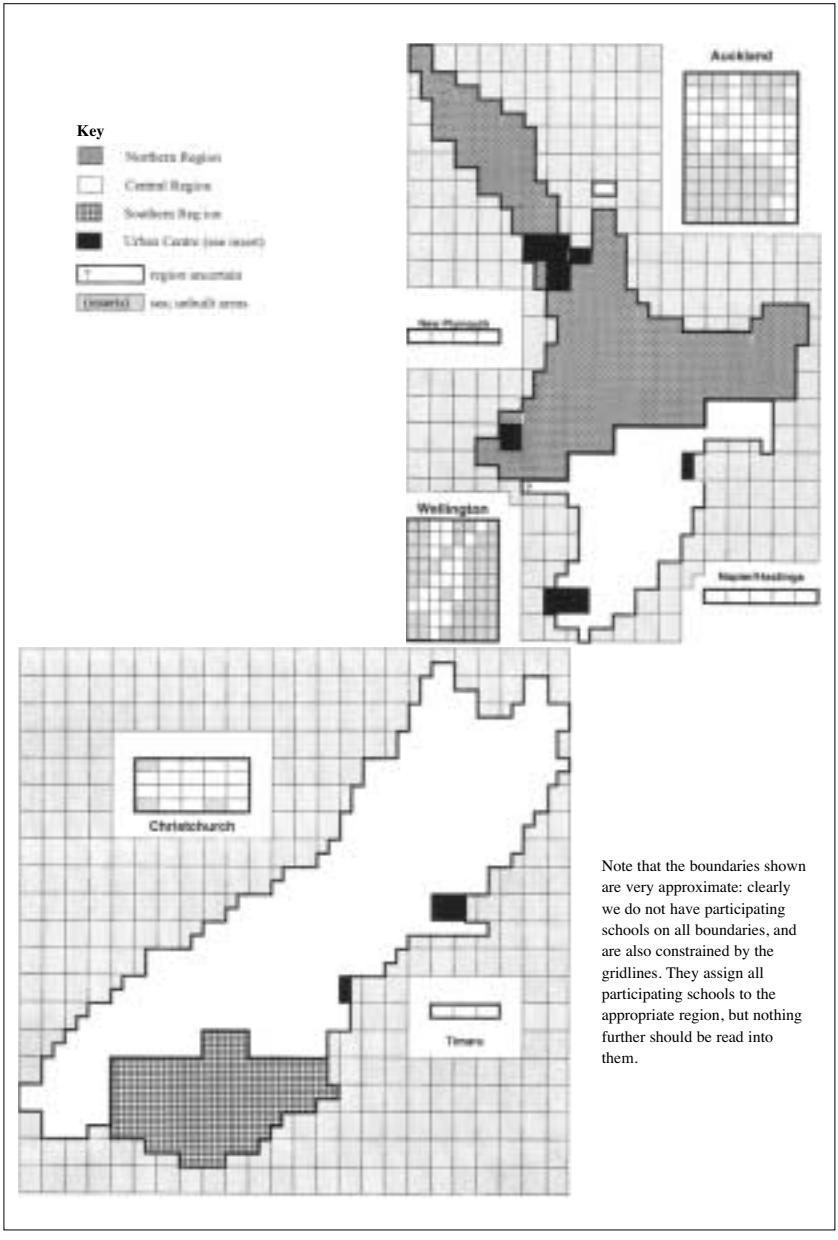
The problems and advantages of the methodology have been discussed in detail elsewhere (e.g. Bauer and Bauer 2000a), and will not be treated further here. However, for the purposes of this paper, it is important to note that we know no details about the individual children who provided specific responses; all we know is the linguistic response and the characteristics and location of the school from which the response came.

The results of the questionnaire as a whole showed that in some sets of data there are three distinct dialect areas in New Zealand, which we call Northern, Central and Southern. The Northern Region extends as far south as Taranaki and the Southern edge of the volcanic plateau. It includes Poverty Bay, but excludes Hawkes Bay. The Central Region extends from Hawkes Bay and the southern fringe of the volcanic plateau across Cook Strait, and down as far as north Otago. It includes the Central Otago lake resorts. The Southern Region consists of East Otago, some of Central Otago, and Southland. Map 1 shows these three regions.

In other cases, the data shows a clear division between the North and South Islands. (There are also a few highly localised forms.) More information about the data which supports these divisions can be found in Bauer and Bauer (2000b).

The data was mapped and graphed to determine which of the many responses showed signs of regionalisation or social differentiation, and then the results for the forms which appeared likely to be interesting were analysed statistically. Full details of the statistical analysis undertaken are not given here, but a brief outline is provided to put the results in context.

Firstly, pairwise comparisons were made between schools to determine the overall level of agreement or difference in their responses. This enabled us to decide on the most likely regional boundaries in those areas where they were unclear from the responses to individual questions. (Taranaki, for instance, sometimes behaved like the Northern Region, and sometimes like the Central Region; comparisons were made between Taranaki schools and all the schools



Map 1: Main Regions

in the areas adjacent to Taranaki, and these showed that Taranaki is more strongly linked to the north linguistically when all the data is taken together.)

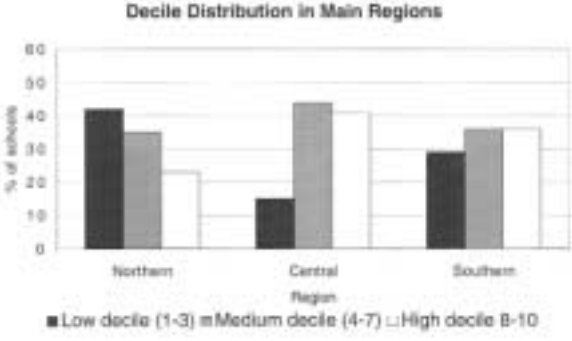
After the regions were determined, each of the responses selected for statistical analysis was analysed in relation to a number of variables (co-variates in statisticians' terms), including Main Region, Island and Decile (see below). A statistical method called Generalized Estimating Equations (GEE) (Liang and Zeger 1986) was used to analyse this data. The statistical package SAS (version 6.12) was used to implement the GEE approach, using PROC GENMOD. This process delivered p-values for each of the co-variates in the analysis in relation to each of the linguistic forms analysed. This method also allowed the interaction between the co-variates to be investigated, so that their relative importance in explaining the variation could be determined.

In the course of analysing the initial data from the questionnaire we formed the hypothesis that the location of the Maori population had a significant effect on the patterns that emerged. In particular, we believe that the distinctness of the Northern Region is often related to the high Maori population in that region, and in some cases, the fact that the North and South Islands differ linguistically can also be attributed to this fact. This paper sets out the evidence which shows this influence of the Maori population on New Zealand dialect areas. In some cases, the connection with the Maori population is direct, since it involves the use of words from the Maori language. In other cases, the vocabulary has been previously established as 'Maori English'. In yet other cases, the vocabulary pertains to concepts which reflect Maori cultural norms. Finally, there are cases where we suggest that it may be possible to deduce a link on the basis of similarity of patterning.

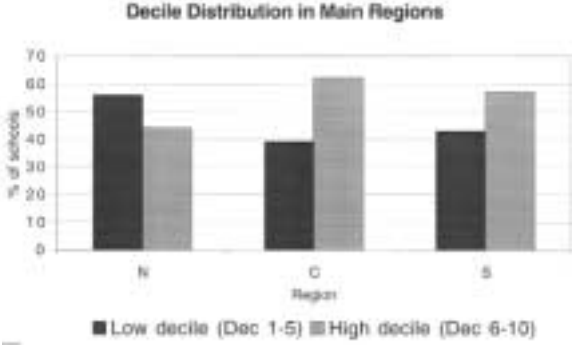
One of the most pervasive effects of the Maori population can be seen in the socio-economic profile of the country. The Ministry of Education gives each school a decile rating based on the socio-economic background of the children attending that school. Schools are put on a socio-economic continuum, and the continuum is divided into ten equal groups, labelled Deciles 1-10. The schools with children from the lowest socio-economic backgrounds are Decile 1 schools, and the schools with children from the highest socio-economic backgrounds are Decile 10 schools. Of course, most schools have children from a variety of socio-economic backgrounds, so the Decile rating of a school reflects only the average socio-economic level of the children in the school.

There is plenty of evidence that Maori students predominate in the lowest socio-economic groups in New Zealand, and that the lowest decile schools

have a lot of Maori children. In 1998, the year before our survey, 42.2% of the children in Decile 1-3 primary schools identified themselves as ethnically Maori, compared with just 6.0% of the children in Decile 8-10 primary schools (Ministry of Education, 1999: 50). From our point of view, what is important is that the decile mix of schools in our sample differs in different parts of the country. In our sample, there are far more schools in the lowest three deciles in the Northern Region, and far more high decile schools in the Central Region, as the following graphs show. The first graph (1a) groups the deciles into three chunks: low decile schools (1-3), medium decile schools (4-7) and high decile schools (8-10). There is a problem with this – there are more deciles in the medium group, which exaggerates the size of this group, but it allows us to isolate the very lowest decile schools. If we divide the deciles into



Graph 1a: Decile distribution of schools in our sample in the three Main Regions

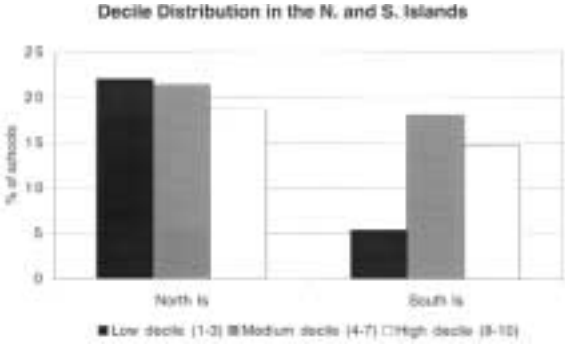


Graph 1b: High and Low Decile schools in our sample in the three Main Regions

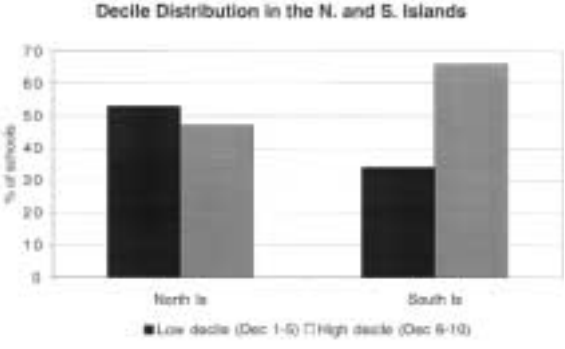
two equal groups, low decile and high decile, the result is shown in Graph 1b. The large Maori population in our Northern Region is an important factor in explaining the preponderance of low decile schools in that region: 65.8% of the Maori children in schools are in our Northern Region (Ministry of Education, 1999: 49) – although note that this figure is for all schools, and not just primary schools.

The two Islands of New Zealand also show the same sort of difference. Graphs 2a and b show that the North Island has almost all the lowest decile schools in our sample, and the South Island has very few low decile schools. Again, the large Maori population in the North Island (compared to the South) is a significant factor in producing the decile imbalances between the Islands.

From the point of view of our data, there are many forms which correlate



Graph 2(a): Decile distribution of schools in our sample in the North and South Islands



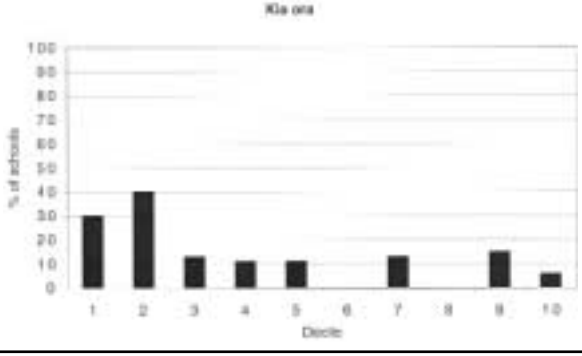
Graph 2(b): Low and High Decile schools in our sample in the two Islands

strongly with both the Northern Region and with low decile schools, or with the North Island and with low decile schools, or with all three of these factors. These factors are often shown by the statistical analysis to be quite closely linked: the forms are Northern because they are low decile and/or they are low decile because they are Northern. What the statistics **do not** show is that in quite a number of cases, these forms are low decile and Northern because that is where the Maori population is found in greatest numbers. We now consider a number of sets of data which show fairly clearly the influence of the Maori population on our regional and social data.

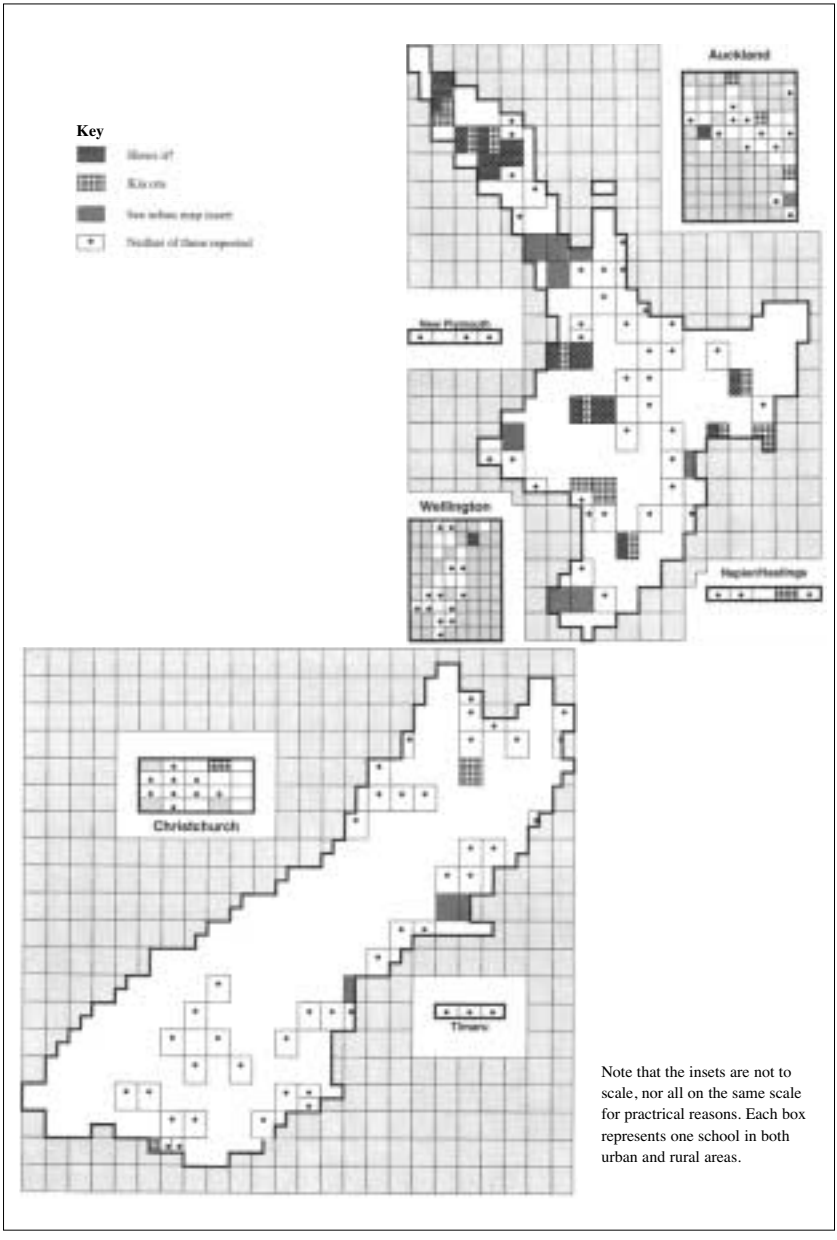
2. Forms overtly derived from *te reo Maori*

There were a few responses to questions which derive directly from *te reo Maori*. (It is impossible to be precise, because in a number of instances, there is doubt: *mucka* probably represents *maka*, but we cannot be sure; however, the best guess is around 30 items.) One of these is the greeting *kia ora*. This greeting was not reported particularly often, but all except three reports were from the North Island, as Map 2 shows.

Note, however, that the statistical analysis showed that the correlation with the North Island rather than the South was not particularly strong (p-value 0.0436). *Kia ora* was also more common in low decile schools, as Graph 3 indicates. The bars on the graph show the percentage of schools in each decile which reported this response. (We do not know whether the reports from higher decile schools came from Maori students in those schools or from



Graph 3: Decile distribution of *Kia ora*



Map 2: *Kia ora* and *Howz it*?

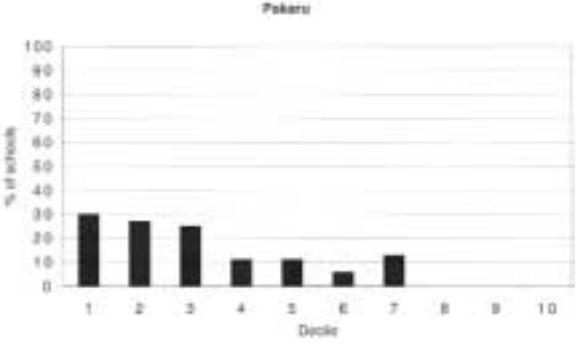
politically correct TP's (teachers' pets) trying to please the teacher. The latter seems likely, since there was some evidence from follow-up visits to 33 schools that Maori words known to Maori children in the high decile schools had not been reported in the original responses.) The p-value for the low decile correlation was 0.0156. The statistics showed that the low decile factor is more important in explaining the distribution of *kia ora* than the North Island factor. In other words, the fact that *kia ora* is a low decile form largely explains why it is more common in the North Island: that is where the low decile schools are.

A second example in this category is *pakaru*, which was elicited in response to the following scenario:

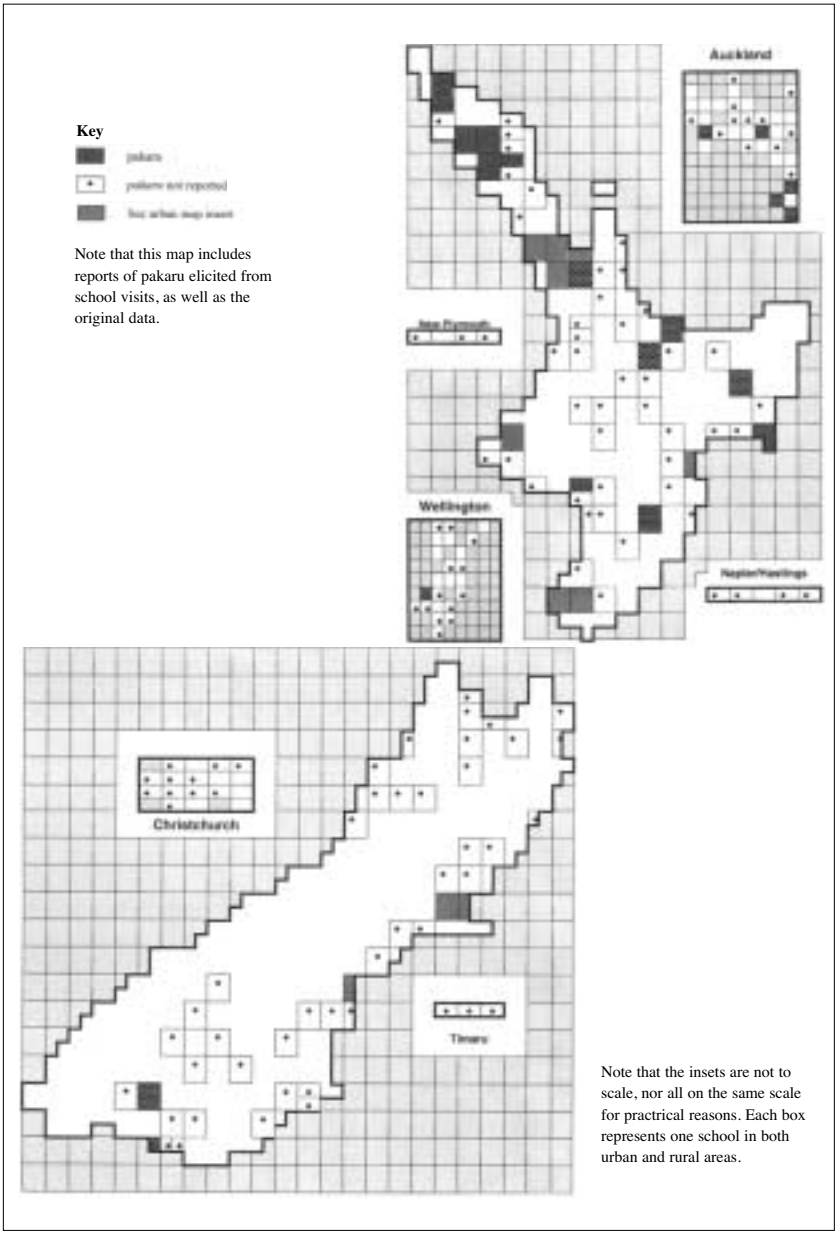
> When you are riding your bike, you lose control, and crash into a bank. Your bike is damaged so badly that you can't ride it. How would you describe the state of your bike?

In the original data there was just one South Island report of this form, but a second school reported using it during the visits to selected schools to interview children (the final phase of the research for the project). Both are shown on Map 3. (Note that the original report came from one of the schools which also reported *kia ora*: it is a decile 2 school; the other is a decile 3 school.)

The correlation with the North Island is still not particularly strong statistically: the p-value is 0.0233. In addition, this form is statistically more common in the Northern Region than the Central Region (p-value 0.0096). It



Graph 4: Decile distribution of *pakaru*


Key
Note that this map includes reports of pakaru elicited from school visits, as well as the original data.
Auckland
Wellington
Christchurch
Timaru
Note that the insets are not to scale, nor all on the same scale for practrical reasons. Each box represents one school in both urban and rural areas.

Map 3: ***Pakaru***

also correlates highly significantly with low decile schools (p-value 0.0002), as Graph 4 shows. (Note that all the reports in Auckland are in South Auckland, for example.) The statistical analysis shows that the fact that this form is Northern is largely due to the fact that it is low decile (and most of the low decile schools are in the Northern Region). The analysis also shows that the fact that this form is more common in the North Island is largely due to the fact that it is low decile (and most of the low decile schools are in the North Island).

3. Forms known to be features of 'Maori English'

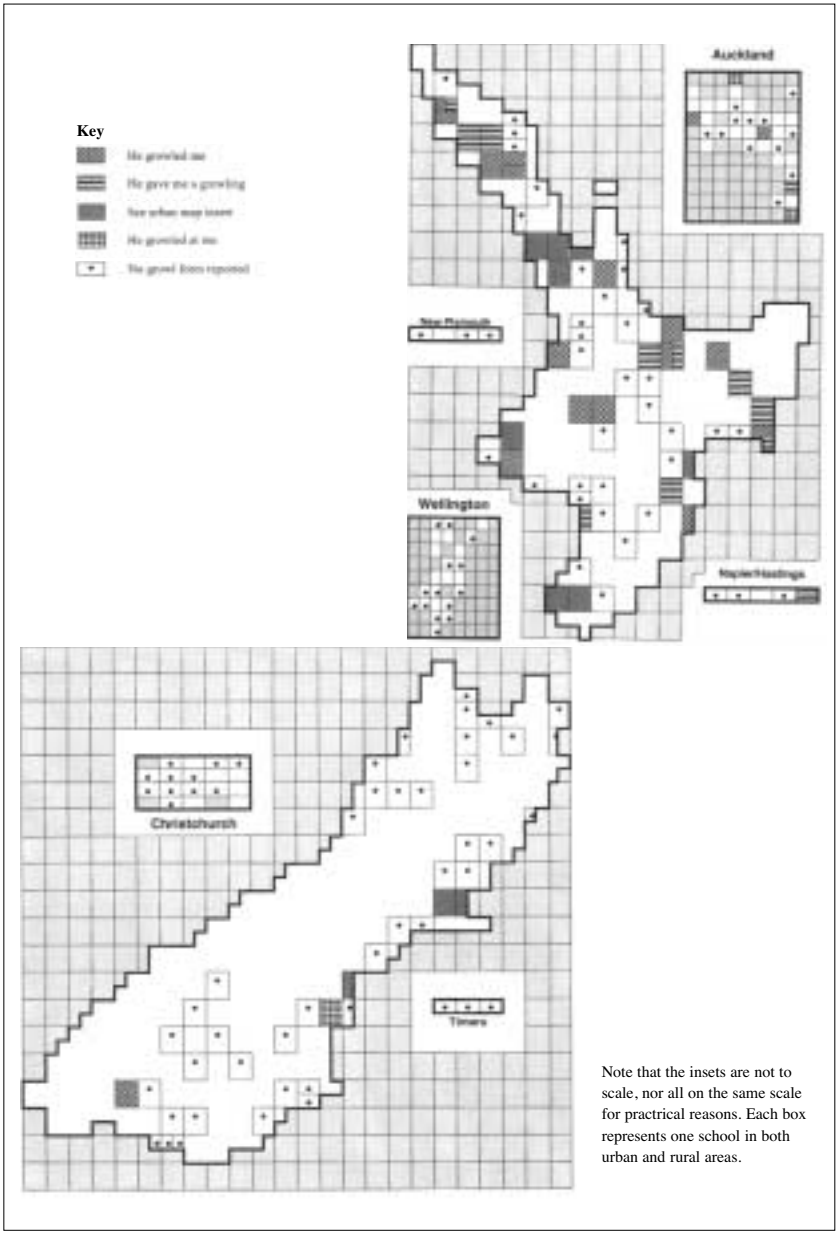
We have put 'Maori English' in scare quotes throughout as a mark of caution. It is not a well-described variety of English, although many – perhaps even most people – believe that they can identify examples of it. Also, while many ethnically Maori speakers use a form of English which is identifiably Maori, not all ethnically Maori speakers do, and in areas with high Maori populations, this form of English is also used by some who are not ethnically Maori.

There are also some items in our data where a term that is known to be a feature of 'Maori English' shows the same kind of distributional pattern as the forms discussed in the previous section. One case is the use of the word *growl*, elicited by the following scenario:

> You ran onto the school garden to get back your ball, and accidentally trod on some plants. The caretaker saw you and told you how cross he was with you. Later, you want to tell your friend what the caretaker did. What would you say?

There were three constructions using the root *growl* in the responses: *he growled me*, *he growled at me*, *he gave me a growling*. *He growled me* is identified by McCallum (1978: 141) as a construction 'which may be unique to Maori speakers of English'. This is a fairly typical case of the transference to English of the Maori construction. (The Maori word for 'growl', *ko(w)hete*, is a transitive verb.) Map 4 shows the distribution of *growl* forms.

Of the two standard English forms, nearly all the responses were *gave me a growling*; there were just two reports of *growled at me*. There are two things to note about the distribution of these forms: the scarcity of any of these in the South Island; and the thinning out of responses in the part of the North Island that falls in the Central Region. (The data obtained from the school visits



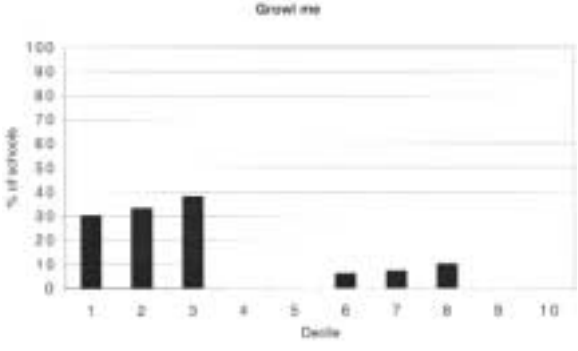
Map 4: *Growl*-forms

showed that in all but two of the South Island schools visited, the children said they would not use *growl* in any form. This included Maori children with North Island connections. However, there were two schools where one child said they would use it, one in the construction *growl me,* and the other in the standard English constructions. The visits thus confirmed the rarity of this in the South Island.)

We did the statistical analysis for both *growl me* and for all *growl* forms taken together. *Growl me* is highly significantly low decile: p-value 0.0006, see Graph 5.

It is highly significantly more common in the Northern Region than the Central Region (p-value 0.0001). In addition, it is just significantly more common in the North Island than the South (p-value 0.0155). (It is also just significantly more common in rural schools than in urban ones 0.0368 – another common correlation with the forms we believe show Maori influence.) The statistics show that the low decile correlation again largely explains the North Island correlation. However, the Northern Region correlation is also very important alongside low decile in accounting for this form. In addition, *growl me* is more strongly associated with low decile in the Northern Region than in the Central Region: the p-values are respectively 0.0168 and non-significant 0.2189.

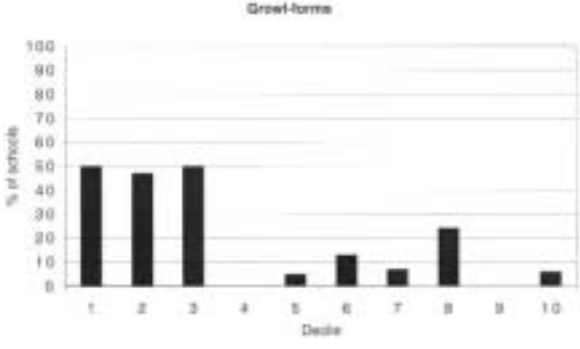
When all the *growl*-forms are considered, they were shown to be highly significantly low decile (p-value 0.0004), see Graph 6. *Growl* forms are also more common in the Northern Region than the Central Region (p-value 0.0001), and more common in the North Island than the South (p-value 0.0011). Here, however, both the regional factors are important alongside



Graph 5: Decile distribution of *growl me*

Decile in accounting for the distribution. Again, *growl*-forms are much more strongly low decile in the Northern Region than in the Central Region: the p-value for the low decile correlation in the Northern Region is 0.0139; the p-value for low decile in the Central Region is not significant (0.2104). (Because there is just one school reporting *growl* in the Southern Region, no sensible statistic is possible there.)

The forms where we have statistics and fairly obvious indications of a Maori connection all show similar sorts of patterns. They are all low decile, and usually also correlate with the Northern Region or the North Island, or both. In all cases, the importance of Decile in accounting for the distribution is shown statistically to outweigh, or at least to equal, the regional factors in importance. There is a strong correlation between the schools reporting these forms and schools in which there is statistically speaking a high Maori population. Thus although we have no way of pinpointing contributions from ethnically Maori children to our questionnaire data, we feel sure that these responses must have been contributed by Maori children or by non-Maori children who have been influenced by the speech patterns of Maori children. While forms like *kia ora* may be taught in classrooms, and their occurrence could reflect this overt teaching, the same is not true of non-standard forms such as *growl me*. While these are produced by speakers who are not ethnically Maori as well as those who are, they appear only where there has been a high degree of exposure to 'Maori English'. Thus they reflect a speech pattern which derives from ethnically Maori speakers, whether they are produced by ethnically Maori speakers or not.



Graph 6: Decile distribution of the root *growl*

4. Forms which pattern similarly to overtly Maori forms

We found some forms which we did not know beforehand to be linked to the Maori population, but where the patterning in terms of decile and regional distribution strongly suggest that this is the case. Three examples are considered.

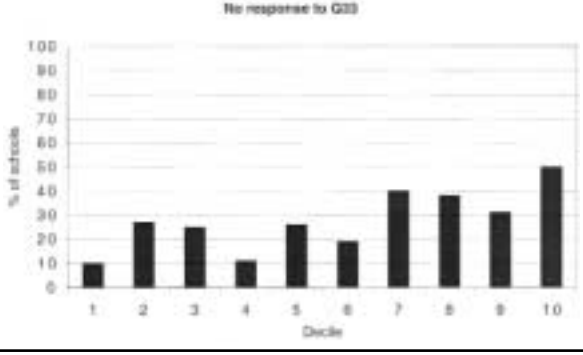
The following question was designed to elicit any words there might be for Maori *whakama* in English. The question was:

> **33** You have just won your school speech competition. The Principal talks to you afterwards and tells you what a wonderful speech it was, and how proud (s)he is of you. You feel very uncomfortable about this. You want to tell your friend **how you felt**. What would you say?

The *whakama* reaction is one of extreme outward embarrassment in the face of praise, even if inside the praise is welcome.

The reactions of the schools to this question were interesting. A large number of schools reported that the children would not feel embarrassed under these circumstances. Some schools even went so far as to comment that it was a stupid question. In other schools, the children responded to the scenario in a way which appears to reflect the *whakama* experience, although they did not use the term *whakama*. The decile distribution of the schools which provided no response to this question is illuminating, see Graph 7. Note the definite tendency for these to be high decile schools.

Two of the responses provided were quite explicit about the dual nature of



Graph 7: Decile distribution of schools providing no response to Q33

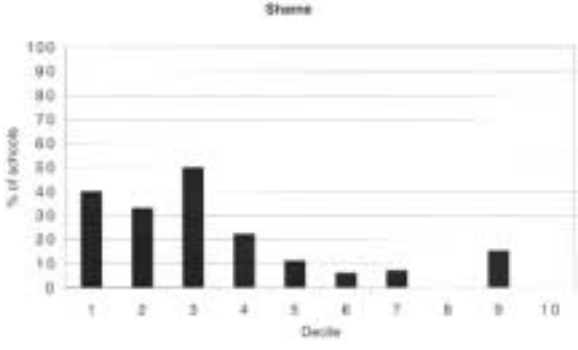
the reaction to this situation: *I felt good on the inside, but not on the outside; a bit shamed but OK*. However, these were both one-off responses. The most common response to this used some form of the root *shame*, most often *I felt shame* or *I felt shamed (out)*. The first of these was frequent enough to show correlations which are of interest in this context.

Shame was shown to be highly significantly low decile (p-value 0.0001), which Graph 8 makes visible.

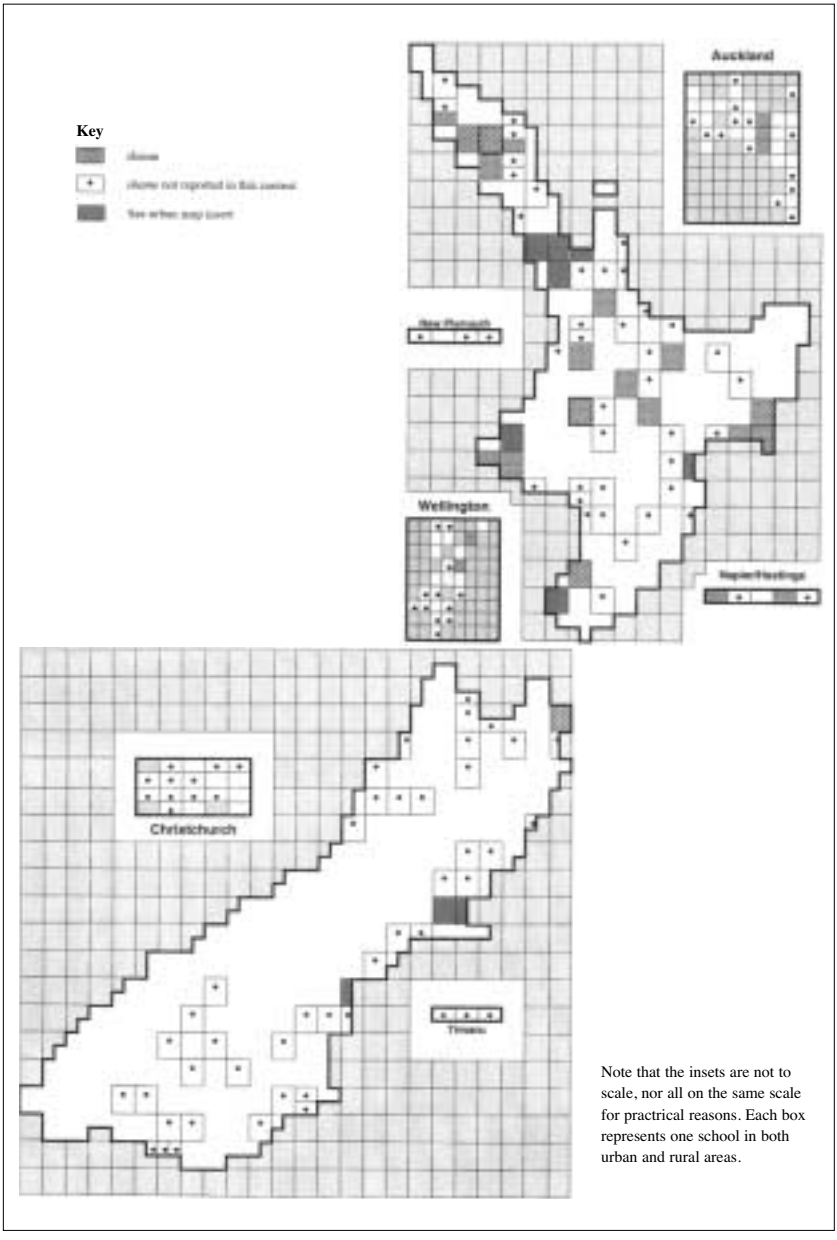
There is also just significantly more use of *shame* in the Northern Region than in the Central Region (p-value 0.0174). It is significantly more common in the North Island than the South (p-value 0.0042), see Map 5.

Decile again is much more important than Island in accounting for this form, although Island is not negligible. Decile also to a very large extent accounts for the difference between the Northern and Central Regions in their use of *shame*. Once again, we see that Decile is the most important factor affecting the distribution of a form which is strongly linked to the Maori population. (While the *whakama* experience may be common to Pacific Island students as well as Maori ones, the majority of the relevant responses here did not come from areas with a particularly high density of Pacific Islands students, 84.7% of whom are in Auckland and Wellington (Ministry of Education 1999: 49)).

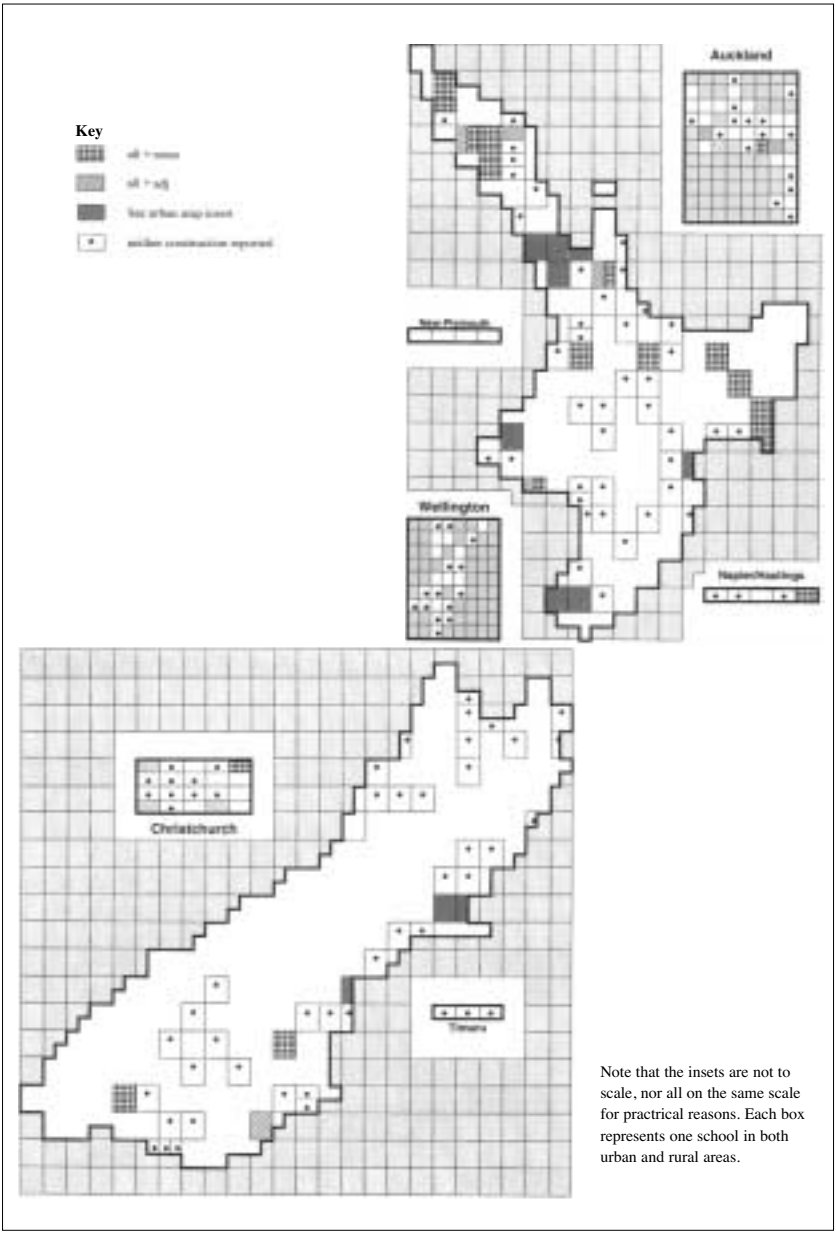
In the answers to several questions, we got responses using the construction *all* + adj or *all* + noun: *It's all good; you're all teko* (Maori for 'wrong', 'lies'); *you're all kaka* (Maori for 'shit'). (We got responses like *bull-kaka*, too: a classic example of code-switching!) Given the occurrence of Maori words in the lexical slot in approximately a quarter of the reports of this



Graph 8: Decile distribution of *shame* in response to Q33



Map 5: *Shame* expressing "whakamaa"



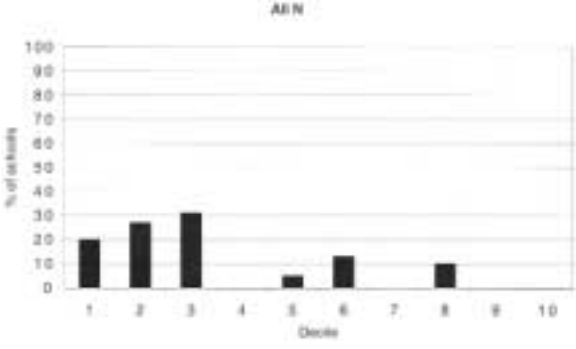
Note that the insets are not to scale, nor all on the same scale for practrical reasons. Each box represents one school in both urban and rural areas.

Map 6: *All* + adj, *all* + noun

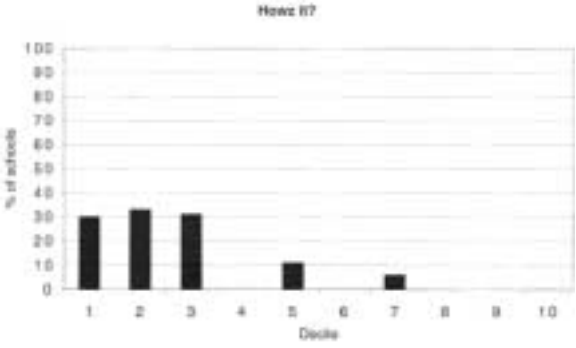
construction, it again seems possible that this construction is produced under the influence of 'Maori English'. The distribution of these forms – from whatever questions they occurred in – is shown in Map 6.

Most of the reports came from schools in the lowest three deciles, as Graph 9 shows for *all* + N. (There were more examples of *all* + N than *all* + Adj, and there were so few of the latter that the graph is not particularly revealing; however it has the same general shape as the graph for *all* + N.) We do not have the statistical analysis for this group of forms, since they were responses to a variety of questions, and were low frequency forms. However, it is highly likely that the statistical analysis would confirm the correlation with low decile, and also confirm the tendency for these to be more common in the Northern Region than the Central Region. The noun forms reported, with number of schools reporting them were: *all (bull)shit* (9); *all crap* (3); *all teko* (3); *all kaka/garks/gacks* (2 – some of these were alternatives from the same school); *all lies* (2); *all plaque* (1); *all class* (1). (The forms *garks* and *gacks* are almost certainly corruptions of *kaka*, and so were grouped with it.) The adjective forms were *all sweet* (2); *all good* (2); *all mushy* (1); *all munted up* (1); *all aggro* (1); *all angus* (1). (There is doubt about the part of speech of the last two, but nothing of consequence hangs on this classification.) There was also one report of *You're all beep*! We take it that the beeped out item was *shit*. (The high decile schools reported *all shit*, *all class*, *all sweet*, *all angus*.) Given the similarity of the patterning of this form and the overtly Maori forms discussed above, we suggest that this may be another previously unidentified construction of 'Maori English'.

A final example is the greeting *Howz it? Howz it?* is highly significantly



Graph 9: Decile distribution of *all* + N

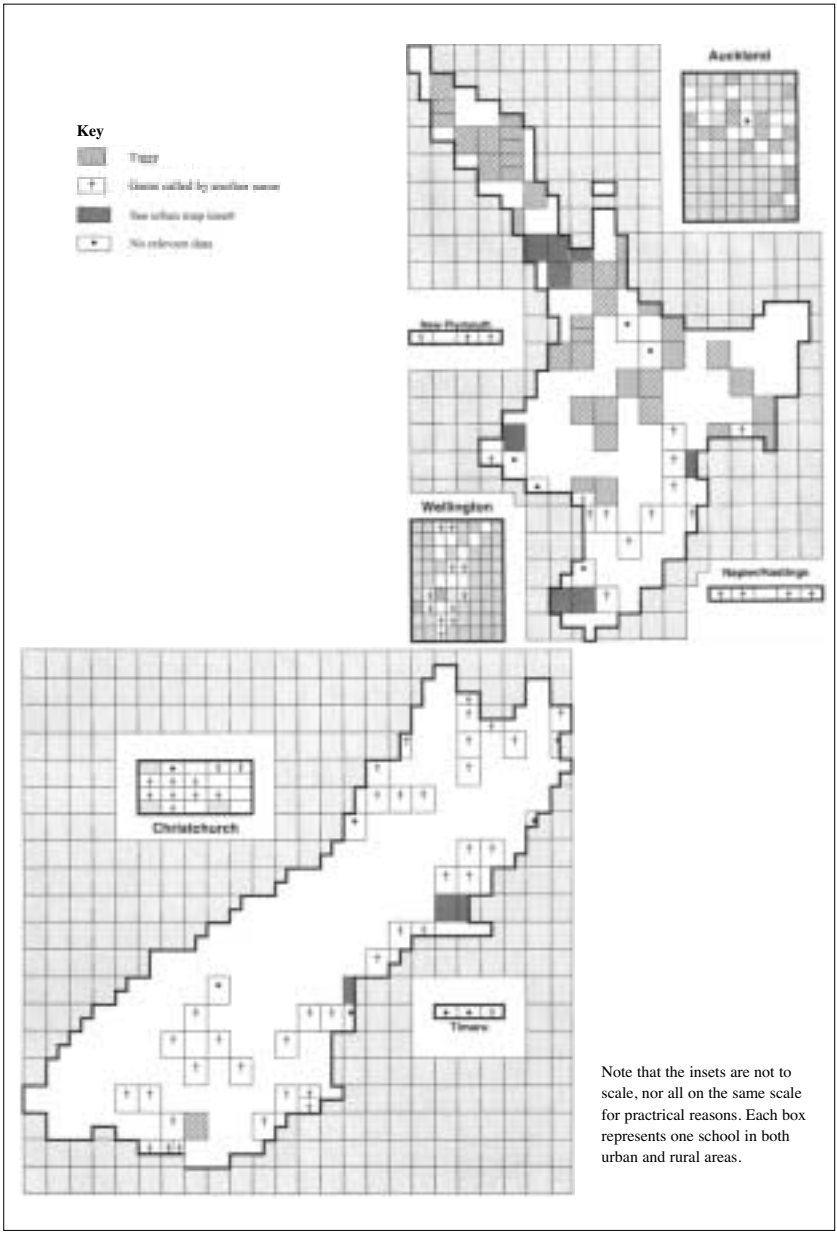


Graph 10: Decile distribution of *Howz it?*

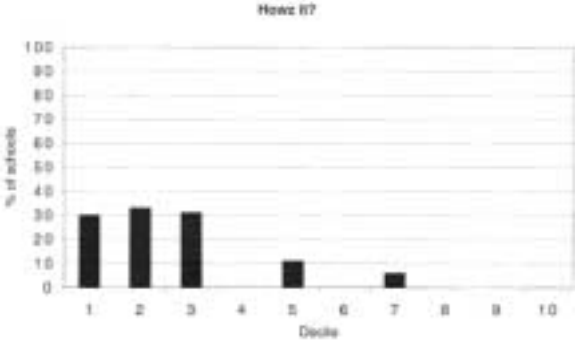
low decile (p-value 0.0004), see Graph 10, was reported only in the North Island, and is significantly more common in the Northern Region than the Central Region (p value 0.0016). The distribution of *Howz it?* is shown on Map 2, alongside the greeting *kia ora*. Many of the schools reporting *Howz it?* are the same schools as reported *kia ora*. Of the ones that did not report *kia ora*, all reported some other form which is typically Maori. The overall patterning of the data is very similar to that seen for other Maori-linked forms: Decile is very important in explaining the distribution of this form, although the fact that it is exclusively a North Island form is clearly also highly significant. This suggests the possibility that *Howz it?* may have been adopted particularly widely in areas where there is a large Maori population. *Howz it?* is not exclusive to New Zealand, but independent of how it is used elsewhere, our evidence would suggest that currently it is yet another characteristic of 'Maori English'.

5. Forms which are not strongly associated with the Maori population

On the other side of the coin, there are forms which are Northern and low decile and not strongly associated with the Maori population. These show rather different characteristics statistically. We consider just one example, the use of *Tiggy* as the name of the chasing game. The distribution of the name *Tiggy* is shown on Map 7, and Graph 11 shows the decile distribution of this name.


Key
Auckland
Wellington
Christchurch
Timaru
Note that the insets are not to scale, nor all on the same scale for practrical reasons. Each box represents one school in both urban and rural areas.

Map 7: ***Tiggy***



Graph 11: Decile distribution of *Tiggy*

Tiggy is a low decile form (p-value 0.0013), it is more common in the Northern Region than the Central Region (p-value 0.0001), and also more common in the Northern Region than the Southern Region (p-value 0.0001). In addition, it is more common in the North Island than the South (p-value 0.0000, derived from a non-zero figure, and so highly significant). However, the statistical analysis for *Tiggy* shows that the regional distribution (i.e. the prevalence in the Northern Region and the North Island) is much more important in accounting for the distribution of *Tiggy* than the decile distribution: *Tiggy* is chiefly a low decile form because it is Northern, and not the other way round. The prevalence in the Northern Region also accounts to a large extent for the fact that this name is more common in the North Island. Thus in forms which we know are not specifically linked to the Maori population, the interaction between Decile and the regional factors is different from those which show a clear link to the Maori population.

6. Conclusions

From the data presented in this paper, we have shown that forms which are closely associated with the Maori population have recurring characteristics in terms of the variables studied. They are associated most strongly with low decile, and also show strong links to either the Northern Region or the North Island or both. It is our hypothesis that other forms which pattern in the same way are likely to be features of 'Maori English' too. Thus this method of gathering data may be a way of identifying other features of 'Maori English'.

It is also clear from the data presented in this paper that Maori people play a significant part in creating the patterns of regional and social differentiation found in New Zealand. In particular, the high density of the Maori population in the Northern areas of the country is one of the most important factors in making the Northern Region so strongly different linguistically from the Central Region, and is often also an important factor in making the North Island linguistically different from the South Island.

Note

1. The research reported in this paper was supported by the Marsden Fund of the Royal Society of New Zealand. We would like to thank all the schools, teachers and children that assisted us in gathering the data for this project. We also gratefully acknowledge the statistical guidance provided by I-Ming Liu of the School of Mathematical and Computing Sciences, Victoria University of Wellington and the comments of two anonymous Te Reo referees.

References

Bauer Laurie and Winifred Bauer. 2000a. 'Two unrelated changes in the English of young New Zealanders.' *Kotare* 3 (1): 3-9.

Bauer Laurie and Winifred Bauer. 2000b. 'Nova Zelandia est Omnis Divisa in Partes Tres.' *New Zealand English Journal* 14: 7-17.

Liang, K. Y. and S. L. Zeger. 1986. 'Longitudinal data analysis using generalized linear models.' *Biometrika* 73: 13-22.

McCallum, Janet. 1978. 'In Search of a Dialect: an exploratory study of the formal speech of some Maori and Pakeha children.' *New Zealand Journal of Educational Studies* 13: 133-143.

Ministry of Education. 1999. *Education Statistics of New Zealand for 1998*. Wellington: Ministry of Education.

ZED TO ZEE REVISITED:

NOTES FROM DUNEDIN ON THE AMERICANISATION OF THE NZE LEXICON THROUGH TIME[1]

James Green: *Psychology Department, University of Otago (PO Box 56, Dunedin)*
Donn Bayard: *Anthropology Department, University of Otago*.
<donn.bayard@stonebow.otago.ac.nz>

Abstract

A sample of 104 Dunedin high school students supplied data on their use of either traditional New Zealand or American vocabulary pairs such as *torch*/ *flashlight* and *lift*/*elevator*. Results are compared with earlier studies carried out in Dunedin in 1984-85 and Auckland 1990. These comparisons suggest that acceptance of American lexical forms in New Zealand English is mediated by both change over time and regional variation.

1. Background

American influence on the lexicon of NZE has been recorded even before the 1840 Treaty of Waitangi. *Creek* in its US sense of a freshwater stream is documented as early as 1815, and *kerosene* for British *paraffin* from 1868 (Orsman 1997: 181, 404). More massive influences began with and after World War II, and with the onset of the 'Pax Americana' of globalisation and US media dominance (Bayard, Weatherall, Gallois, and Pittam 2001) the steady stream (or *creek*, if you will) has become a flood. There have been a number of studies devoted to the prevalence of incoming American forms in NZE (Bayard 1987, 1989, Meyerhoff 1993, Vine 1995, 1999), but only one

study has attempted to examine the question of longitudinal change through time. In 1995 Leek and Bayard published a comparative study of US versus 'traditional' NZE vocabulary among two fairly large and well-stratified bodies of informants in Dunedin (1984-85; 144 informants) and Auckland (1990; 300 informants). The study asked for informants' actual use of one or the other of pairs like *lift-elevator* and *torch-flashlight*, and then asked which of the pair was 'better English'. The lists of words used in the two studies were not identical, but 18 pairs were shared between the two studies.

We thus had studies carried out at opposite ends of New Zealand and separated in time by five to six years. In 1995 Leek and Bayard asked two questions; one of these received a fairly firm answer, but the other did not. It did indeed appear that preference for one of a pair as 'better English' was a good predictor of its being accepted for usage over the 'traditional' alternative. But our second question, on the marked preference for American alternatives in Auckland as compared with Dunedin, was unanswerable with the data to hand. Was the greater Auckland use and preference of incoming Americanisms due to the fact that the survey was carried out five or six years later than the Dunedin one, or because Auckland as New Zealand's 'Queen City' is simply more cosmopolitan and exposed to world trends than relatively isolated Dunedin? We now have a small body of data from Dunedin, collected 15-16 years after the Dunedin survey and 10 years after the Auckland survey, which sheds some light on these two questions.

2. The 2000 Dunedin high school sample

One hundred and four Dunedin High School students participated as part of an education expo for the tertiary sector, held at Dunedin Stadium in September 2000. Participation was voluntary, simply by filling out a one-page questionnaire setting out each pair of terms in context, and asking the students to circle the word they 'normally say'. It was also self-selected, in that students chose to fill out the questionnaires at the expo. Only questionnaires that listed the student's school were processed to avoid including a few interested adults in the sample. Ages were not asked for, but it is safe to assume that the students were in Years 9-13 (Form III-Form VII), or about 13 to 18 years old.

3. Results

The results from the current questionnaire are presented in Table 1, along with comparative data drawn from the Dunedin survey carried out during 1984-5 (Bayard 1989) and Auckland in 1990 (Leek and Bayard 1995). As both of these earlier surveys used stratified samples covering the entire age spectrum, sub-samples were used; these were 48 informants from Dunedin aged between 12 and 25, and 60 from Auckland in the narrower age range of 15 to 18. Thus the age ranges of the three samples do not correspond exactly, but are certainly close enough to be as generally satisfactory as other comparative studies on features of NZE (e.g., Batterham 2000, Allan and Starks 2000).

For the thirteen items common to the three datasets the mean innovative use in Dunedin in 1984-5 was 10.4%, in Auckland in 1990 18.2%, and a decade later in Dunedin in 2000 mean innovative use was 20.2%. It would thus appear that Dunedin is perhaps 10 years behind Auckland in the use of American innovations.

This contention can be further examined by looking at different possible patterns of usage present in the sample. If change was due solely to the progress of time then the following chronological pattern would be expected:

Dunedin 1984-5 —> Auckland 1990 —> Dunedin 2000

The change from *pictures* to *movies* is the only one strongly fitting this pattern, with *biscuit* to *cookie* and *rubber* to *eraser* showing some support. *Bonnet* to *hood* shows a non-significant trend supporting this.[2]

The second possibility, suggested by the mean use of innovation already reported, is that Dunedin is around 10 years behind in the use of American innovations, as represented in the following pattern:

Dunedin 1984-5 —> Auckland 1990 = Dunedin 2000

There is little evidence for this, although *bonnet* to *hood* could be interpreted as such.

A third possibility is that Dunedin is somewhat immune to the use of American innovation and that Auckland shows the greater use of innovations while Dunedin has stayed constant over time:

Dunedin 1984-5 = Dunedin 2000 —> Auckland 1990

	DUNEDIN (1984-5): USE	AUCKLAND (1990): USE	DUNEDIN (2000): USE	DUNEDIN (1984-5): PREF.	AUCKLAND (1990): PREF.
Torch vs. FLASHLIGHT	0	8	3	33	57
Lift vs. ELEVATOR	8	22	16	63	93
Rubber vs. ERASER	2_a*	7_a	18_b	52	87
Flats vs. APARTMENTS	—	17b	5a	—	68
Zed vs. ZEE	13_a	—	32_b	—	—
Frock vs. DRESS	—	100	100	—	80
Jersey vs. SWEATER†	8	13	11	15	43
Boot vs. TRUNK	0	3	6	33	27
Note vs. BILL	11_b	20_b	3_a	21	20
Tinned vs. CANNED Food	—	—	59	—	—
Biscuit vs. COOKIE	2_a	5_{ab}	13_b	2	2
Petrol vs. GAS	0_a	22_b	7_a	6	3
pictures vs. MOVIES	36_a	82_b	100_c	38	65
Nappies vs. DIAPERS	4	3	4	27	47
Bonnet vs. HOOD	16	28	32	10	27
Footpath vs. SIDEWALK	4	2	7	17	23
Serviette vs. NAPKIN	44_b	22_a	43_b	31	30
Fizzy Drink vs. SODA POP	—	—	3	—	—
Ice Blocks vs. POPSICLES	-—	—	21	—	—
Lollies vs. CANDY	—	—	5	—	—

* Percentages not sharing the same subscript differ at p=0.05 using the chi-square statistic calculated on raw data; ‘a’ indicates the least frequent use, ‘b’ second least frequent, and ‘c’ most frequent use of the American innovation.

† The Auckland contrast was between SWEATER and JUMPER.

Table 1: Percentage of students using American innovation by sample (Dunedin 1984-5: 48 aged between 12 and 25; Auckland 1990: 60 aged between 15 and 18; Dunedin 200: 104 presumably aged between 13 and 18)

This trend is fitted by a number of changes – *flat* to *apartment*, *note* to *bill*, and *petrol* to *gas*. *Note* to *bill* even appears to show reduced usage in Dunedin across time. *Torch* to *flashlight* also shows this pattern of usage. *Serviette* to *napkin* contradicts this, however, with higher use of *napkin* in Dunedin at both time points.

Finally, exploratory analyses on the Dunedin 2000 sample using a crude measure of socioeconomic status (school decile ratings) revealed no significant pattern of correlations when controlling for between-school differences. The results instead suggested that between-school differences, that is individual school 'cultures', may have a stronger influence than the socioeconomic background of the schools.

4. Discussion

Although the limited data from the present survey make in-depth discussion unwise, integrating the results of the present study with previous research conducted by Bayard and Leek suggests that a number of trends are possible. Evidently the discrepancies between their two previous studies — Dunedin in 1984-5 and Auckland in 1990 — may be based in both regional and temporal differences.

Use of American innovations is generally less in Dunedin than in Auckland, and in some cases it seems that Dunedin is keeping to the original New Zealand usage more than Auckland; where change is occurring in Dunedin, this is happening later than in Auckland. There are exceptions to this rule, and it is possible that these may be explained by sampling differences. The present study contains only teenagers, whereas the Dunedin subsample used here includes informants in their early 20s (mean age of the 48 is 20.3). That *movies* (relative to *pictures*) was universally preferred in the present study may be a function of the younger demographic. This could potentially explain the increased use of *hood* (for *bonnet*) in the present study. The rising use of *cookie* (in relation to *biscuit*) may be in part due to the start of production, and exceptional popularity, of the 'Cookie Time' biscuit. The high use of *napkin* in Dunedin at both time points would perhaps suggest, not necessarily American influence, but the existence of some lexical differences by region (as with words for *tag*; Bauer and Bauer 2000). Further investigation of regional variation could include the genesis of the new meaning of *sifting*, and novel terms such as *mare*[3] in Dunedin's student culture.

In 1989 Bayard discussed the possibility of 'change from above' and 'change from below' occurring in vocabulary along the lines originally proposed by Labov for sound changes (Bayard 1989: 32-33). These were defined as changes possibly viewed as prestigious coming in more or less consciously,[4] with preference figures notably higher than use figures (*elevator*, *flashlight*, *eraser*); and ones apparently being introduced less consciously and with lower prestige, as evidenced by use percentages being much higher vis-à-vis preference figures (*gas*, *dollar bill*, *sweater*). We can see examples of these two general patterns in words like *elevator*, *eraser*, *movies*, *trunk*, and to some extent *sidewalk*; all would appear to be changes from above, with greater preference figures in 1985-85 and increased use figures in 2000. As possible examples of change from below we have *cookie*, *hood*, *napkin*, and *sweater*. It is puzzling that use figures for *gas*, *dollar bill*, and *flashlight* have actually declined in the interval between the two Dunedin studies. The Auckland data published in 1995 allowed us to postulate that a general relationship was present between greater preference and a later increase in use figures, and the above examples from our present data seem to support this relationship; but note that there is of course no 'precise predictive ratio between changes in preference and usage' (Leek and Bayard 1995: 120).

The motivations behind such lexical shift are of course another matter, and one which was addressed in Vine's recent study (1999). Her data is derived from a small sample of 30 female informants, of whom only ten aged 20-29 approach the age range of the present study.[5] Vine feels that Leek and Bayard's 1995 study implies 'that people will only adopt American terms if they are unaware of the term's origins' (1999: 13). Her research rather suggests that 'speakers are frequently oblivious to the origins of terms' (loc. cit.). This is in fact precisely the point Leek and Bayard make (1995: 123); in Bayard's experience, informants were often unaware of which item of a pair was British and which American, even when a dislike of 'Americanisms' was expressed. There is no doubt about a covert fondness for the North American accent, as documented in almost all accent evaluations carried out in New Zealand (Bayard 2000; Bayard, Weatherall, Gallois, and Pittam 2001; Vornik 1999). However, the whole question of motivation for the changes discussed here is beyond the scope of this paper, and clearly requires further investigation.

Finally, further research should perhaps focus on simultaneous sampling in different regions around New Zealand, enabling more accurate distinction between temporal and regional variation. This brief study suggests both factors are important.

Notes

1 Green carried out the coding and the analysis of the questionnaires; Bayard devised the list and is responsible for the more general sections of the paper. We both thank Catherine Waite, of the Otago Anthropology Department, for suggesting the study, and for designing and collecting the questionnaire.
2 Significance was determined using the chi-square statistic calculated on the raw data.
3 *Sifting* is aimless or subversive wandering; *mare* is an unfortunate or undesirable event.
4 In at least some cases (*elevator*, *eraser*, *apartment*) the Latinate appearance of the word may well have an influence.
5 Similarly, Meyerhoff's 1993 study is based on a sample of 60 working-class Māori and Pākehā (NZ European) informants, but only 20 were at all comparable with the three samples discussed here, and they were older (20-29) (Meyerhoff 1993: 235).

References

Allan, W. Scott, and Donna Starks. 2000. '"No-one sounds like us?" A comparison of New Zealand and other southern hemisphere Englishes.' In Allan Bell and Koenraad Kuiper (eds) *New Zealand English*. Amsterdam/Wellington: John Benjamins/Victoria University Press. 53–83.

Batterham, Margaret. 2000. 'The apparent merger of the front centring diphthongs — EAR and AIR — in New Zealand English.' In Allan Bell and Koenraad Kuiper (eds) *New Zealand English*. Amsterdam/Wellington: John Benjamins/Victoria University Press. 111–145.

Bauer, Laurie and Winifred Bauer. 2000. 'Nova Zelandia est omnis divisa in partes tres'. *New Zealand English Journal* 14: 7-17.

Bayard, Donn. 1987. 'Class and change in New Zealand English: A summary report.' *Te Reo* 30: 3-36.

Bayard, Donn. 1989. '"Me say that? No way!": the social correlates of American lexical diffusion in New Zealand English.' *Te Reo* 32: 17-60.

Bayard, Donn. 2000. 'The cultural cringe revisited: Changes through time in Kiwi attitudes toward accents'. In Allan Bell and Koenraad Kuiper (eds) *New Zealand English*. Amsterdam/Wellington: John Benjamins/Victoria University Press. 297–322.

Bayard, Donn, Ann Weatherall, Cynthia Gallois, and Jeffery Pittam. 2001. 'Pax Americana? Accent attitudinal evaluations in New Zealand, Australia, and America.' *Journal of Sociolinguistics* 5 (1): 22-49.

Leek, Robert-H. and Donn Bayard. 1995. 'Yankisms in Kiwiland, from zed to zee: American lexical and pronunciation incursions in Dunedin (1984-1985) and Auckland (1990)'. *Te Reo* 38: 105-125.

Meyerhoff, Miriam. 1993. 'Lexical shift in working class New Zealand English: variation in the use of lexical pairs'. *English World-Wide* 14 (2): 231-248

Orsman, Harry. 1997. *The Dictionary of New Zealand English: A Dictionary of New Zealandisms on Historical Principles*. Auckland: Oxford University Press.

Vine, Bernadette. 1995. 'American English and Wanganui women's speech.' *New Zealand English Newsletter* 9: 23-26.

Vine, Bernadette. 1999. 'Americanisms in the New Zealand English lexicon.' *World Englishes* 18: 13-22.

Vornik, Lana A. 1999. 'The influence of accent on the misinformation effect.' Unpublished MSc thesis, Department of Psychology. Wellington, N.Z.: Victoria University of Wellington.

NEW ZEALAND BOOKS AND THESES 2000–2001

Compiled by John Newman: *School of Language Studies, Massey University. (Private Bag 11 222 Palmerston North) <j.newman@massey.ac.nz>*

Books

2000

Bell, Allan and Koenraad Kuiper. (eds) 2000. *New Zealand English*. Wellington: VUW Press and Amsterdam: Benjamins.

Crowley, Terry. 2000. *An Erromangan (Sye) dictionary*. Canberra: Pacific Linguistics.

Holmes, Janet . 2000. (ed) *Gendered Speech in Social Context: Perspectives from Gown and Town*. Wellington: Victoria University Press.

Kearns, Kate. 2000. *Semantics*. Basingstoke, Hampshire: Macmillan and New York: St. Martins Press.

Li, FuYin and Koenraad Kuiper (eds) 1999. *Semantics: A Course Book*. Shanghai: Shanghai Foreign Languages Education Press.

Song, Jae Jung. 2001. *Linguistic Typology: Morphology and Syntax*. Harlow and London: Pearson Education.

2001

Barnard, Roger and Ray Harlow (eds) 2001. *Proceedings of the Conference 'Bilingualism at the Ends of the Earth'*, University of Waikato, November 2000. Hamilton: Dept of General and Applied Linguistics, University of Waikato.

Bauer, Laurie 2001. *Morphological Productivity*. (Cambridge Studies in Linguistics 95) Cambridge: Cambridge University Press.

Harlow, Ray. 2001. *A Maori Reference Grammar*. Auckland: Pearson Education.
Lynch, John and Terry Crowley. 2001. *Languages of Vanuatu: A New Survey and Bibliography*. Canberra: Pacific Linguistics.
Lynch, John, Malcolm Ross and Terry Crowley. 2001. *The Oceanic Languages*. London: Curzon Press.
Song, Jae Jung. 2001. *Toward a Typology of Causative Constructions*. (Languages of the World No. 23) Mūnchen: Lincom Europa.
Wouk, Fay and Malcolm Ross (eds) 2001. *The Historical and Typological Development of Western Austronesian Voice Systems*. Canberra: Pacific Linguistics.

Theses

2000

Borgen, Anette. 2000. High Rising Terminals in New Zealand. BA (Hons.) thesis. The University of Auckland.
Kim, Jeong. 2000. Language diaries: an investigation into language use in the New Zealand Korean community. MA thesis. University of Auckland.
Kochovska, Slavica. 2000. Palatalization processes in standard Macedonian: A non-linear approach. MA thesis. The University of Auckland.
Lo, Josephine. 2000. A syntactic analysis of Cantonese verb sequencing. MA thesis. The University of Auckland.
Nakanishi, Nobuko. 2000. Language maintenance and language shift in the Japanese community of Auckland: A study of the interaction between the sojourners and the immigrants. MA thesis. The University of Auckland.
Owens, Melanie. 2000. Agreement in Bimanese. MA thesis. University of Canterbury.
Setiawan, Slamet. 2000. Language shift in a bilingual community: the case of Javanese shift in Surabaya, Indonesia. MA thesis. The University of Auckland.
Sumbuk, Kenneth Memson. 2000. Morphosyntax of Sare. PhD thesis. Waikato University.
Tent, Jan. 2000. The dynamics of Fiji English: a study of its use, users and features. PhD thesis. University of Otago.
Torii, Shizuka. 2000. 'Wa'/'ga' subjects in Japanese and subdivisions of tense. PhD thesis. Victoria University of Wellington.

2001

Aitchison, Therese. 2001. An HPSG analysis of the Tongan definitive accent. MA thesis. Canterbury University.

Brown, T. Pascal. 2001, How do people get things done at work? A discourse analysis of directives in an Auckland factory. MA thesis. Victoria University of Wellington.

de las Cuevas Torresano, Felix. 2001. The acquisition of the subjunctive in Spanish as a second language. MA thesis. Victoria University of Wellington.

Halliday, George. 2001. Stative sentences in Japanese and the role of the nominative marker ga. MA thesis, Massey University.

Hodge, Mellanie J. 2001. /ju:z/ are in the paddock eating grass: A study of the acceptance of grammatical variants among adolescents from Southland and Otago. MA thesis. University of Otago.

Hunkin, Galumalemana Alfred. 2001. A corpus of contemporary colloquial Samoan. MA thesis. Victoria University of Wellington.

Innes, Bronwen. 2001. Speaking up in court. PhD thesis. The University of Auckland

Musgrave, Jill. 2001. A grammar of Nefe'ei, Vanuatu. MPhil thesis. University of Waikato.

Oshika, Yuko. 2001. Intercultural conflicts between New Zealand and Japanese under business conditions. MA thesis. Massey University.

Park Seon-Ho. 2000. A sociolinguistic and social psychological study of Korean communities in New Zealand: language maintenance and second language learning. PhD thesis. Victoria University of Wellington.

Sunahara, Takako. 2001. Interpersonal communication: effective communication with Japanese people. MA thesis. Massey University.

Vine, Bernadette. 2001. Workplace language and power: directives, requests and advice. PhD thesis. Victoria University of Wellington.

PERSONALITY IN THE SOCIOLINGUISTIC INTERVIEW SITUATION

Stacey Nicholas: *Department of Linguistics, University of Canterbury (Private Bag 4800, Christchurch, NZ)*

Abstract

Personality as an influential factor in the sociolinguistic interview situation is a topic which has been the focus of much speculation, but little research seems to have been done by sociolinguists. This paper presents the findings of a study into aspects of personality in the sociolinguistic interview situation. Gregariousness is the facet of extroversion with which people most commonly associate the notion of an 'extrovert'. Although this facet was expected to have a significant effect on the participants' percentage of talk time, the relationship was found to be non-significant. However there was a significant relationship between another facet of Extroversion, namely Warmth, and the percentage of time the participants spoke in the interview. Significant results were unexpectedly obtained for the association between duration of interviewer speech and duration of interviewee speech. Clear patterns were also found between the way interviewees said 'No' to questions asked by the interviewer, and personality.

1. Introduction

This paper presents the results of a study on personality as a significant factor in the sociolinguistic interview situation. Although sociolinguists do not seem to have studied the effects of personality on the interview situation, an analysis of the interview situation in terms of two-person interaction can be found in the literature of psychology. The present study investigated the relationship between amount of speech and personality facets as identified by the NEO PI-

R, a standard psychology questionnaire used to measure personality. I hypothesised that amount of talk in the interview would be significantly influenced by the personality of the informant.

2. Background

2.1 Personality Type

The NEO PI-R personality questionnaire (Costa and McCrae 1992) was used to measure personality in this study. The NEO PI-R is a revised form of the NEO Personality Inventory, and is based on the 'Big Five' approach to personality psychology, a taxonomy of personality traits developed over a period of time by researchers such as Allport and Odbert (1936), and Cattell (1946). Five factors of personality make up the model, commonly remembered by the mnemonic OCEAN: Openness, Conscientiousness, Extroversion, Agreeableness, and Neuroticism.

Costa and McCrae's NEO PI-R personality questionnaire is trait-based, and employs the Big Five (i.e., OCEAN) as domains. McCrae and Costa (1990: 177) argue that traits from the five-factor model of personality can be measured 'with an acceptable degree of accuracy by either self-reports or ratings from knowledgeable sources'. Their studies also demonstrate 'that over the adult portion of the life course there is little change in the average level of most commonly measured personality traits' (McCrae and Costa 1990: 177). The NEO PI shows cross-cultural stability, and, as noted by Pervin 'there is growing evidence that people in diverse cultures, using very different languages, construe personality in accord with the five-factor model' (1993: 308–309).

The NEO PI-R questionnaire consists of 240 questions, 48 for each of the five domains. The response to each question is made on a Likert scale, a five-point scale ranging from 'Strongly Agree' to 'Strongly Disagree'. Each domain consists of six facets, each of which is assessed by eight questions. Facets are more specific traits, which, when formed into a cluster, constitute a domain. The questionnaire gives a set of scores for each facet and an overall score for each of the domains.

After the NEO PI-R professional manual was consulted, the two domains which were hypothesised to be most relevant to this study were Openness and Extroversion. Due to time constraints and the difficulties of drawing on a dataset many times in statistical analysis, it was not possible to assess the

influence of all five domains and their facets on amount of speech in the sociolinguistic interview. Brief descriptions of the Openness and Extroversion domains and their facets, based on the NEO PI-R professional manual (Costa and McCrae 1992), are provided below.

Facets of Openness are designated by the *aspect* or *area of experience* to which the person is open. Facets under the domain 'Openness to experience' are Fantasy, Aesthetics, Feelings, Actions, Ideas, and Values. Individuals who are open to Fantasy have a vivid imagination and an active fantasy life. High scorers on the Aesthetics scale have a deep appreciation for art and beauty. Those people who score highly on the Feelings facet experience deeper and more differential emotional states and feel both happiness and unhappiness more intensely than others. Openness to Action is realised behaviourally in the willingness to try different activities, go to new places, or eat unusual foods. The Ideas facet is characterised by open-mindedness and a willingness to consider new, perhaps unconventional ideas. Openness to Values means the readiness to re-examine social, political, and religious values.

The Extroversion domain includes facets of Warmth, Gregariousness, Assertiveness, Activity, Excitement-seeking, and Positive emotions. Introversion may be realised as the absence of Extroversion rather than being the opposite of Extroversion. Warmth is the facet of Extroversion most relevant to interpersonal intimacy issues—Warm people are affectionate and friendly and genuinely like people. Gregariousness is the preference for other people's company, and is the facet which most people think of, when they talk about Extroverts. High scorers of the Assertiveness facet are dominant, forceful, and socially ascendant. A high Activity scorer displays a need to keep busy, to lead a fast-paced life. High scorers on the scale of Excitement-seeking crave excitement and stimulation, and like bright colours and noisy environments. The facet of Positive emotions is the tendency to experience positive emotions such as joy, happiness, love and excitement.

Two facets of Extroversion and two of Openness were used in this study. The reason why specific facets—rather than whole domains—were investigated, was that thorough research into the descriptions in the NEO PI-R professional manual led to the conclusion that not every facet of Extroversion and Openness was pertinent to the specific situation of a two-person speech interaction. The four facets chosen were: Warmth, Gregariousness (Extroversion), and Ideas and Fantasy (Openness). It was hypothesised that these facets would be most significant in influencing quantity of speech in the interview situation. The facets of Ideas and Fantasy were chosen because of

the 'Openness to experience' nature of the questions used in the interview (see *'The Interview'* in Methodology section). That is to say, the Openness to Experience facets of Ideas and Fantasy relate to the questions dealing with ghosts/UFO's and danger of death, as openmindedness of the paranormal or the supernatural may have influenced participants in this study into answering the way they did. For example, one could imagine that if a participant was not openminded about these kinds of phenomena, then it could follow that the participant may be curt or not interested in answering the question. Gregariousness was chosen because it is the facet of Extroversion which corresponds with the notion of an 'Extrovert' in the general sense. Warmth was also investigated as it is the facet of Extroversion most related to issues of interpersonal intimacy, and it therefore seemed that it would be the facet most likely to relate to conversational interaction between people.

3. Relevance of personality in the interview situation

3.1 Sociolinguistic considerations [1]

Few sociolinguists have speculated on the relevance of personality in the interview situation. Past sociolinguistic studies looking at the interview situation have neglected the possibility of personality influencing the interview situation, and have instead looked at the effects of addressee status or solidarity, gender, insider versus outsider status and ethnicity [2] (see Rickford and McNair-Knox 1994: 236).

Wolfram and Fasold acknowledge that personality is a component of the sociolinguistic interview, but believe that it cannot be controlled for (1974: 54). Other linguists acknowledge personality as an integral part of situation. Brown and Fraser (1979: 56) state that 'it is clear that situational factors, both participant and nonparticipant ones, are interlinked in highly complex ways: class is related to power and status at an interpersonal as well as institutional level, and mood, personality, social relationship, purpose and setting are all related.' Robinson (1972: 144) notes that 'Two separable theoretical issues are necessarily linked in natural situations—personality and role relationship; we expect role relationships to constrain verbal behaviour, but personality characteristics also affect what is said.'

Other authors discuss the topic of verbal output and personality (Scherer 1979: 118; Scherer and Giles 1979: 178). Scherer and Giles (1979: 178) have reviewed the literature and remark that 'extroversion seems to be the only trait

which is consistently found to be associated with a greater amount of verbal output or longer total speaking time.' The problem here is that 'extroversion' is not defined and there are potentially many ways to define it. Scherer and Scherer (cited in Scherer 1979: 119) have argued that personality traits and attitudes are more likely to determine behaviour than are situational factors.

Furnham (1990: 77–78) claims that there are six possible relationships between personality and speech/language. The first possibility is that there is no such relationship; the second is that personality 'determines' speech; and the remaining options deal with the possibility that speech 'determines' personality, that personality and speech are reciprocally determined, that there are mixed relationships between personality and speech, and finally, that personality and speech are moderated by other variables. This paper is concerned with the concept of personality being one of the factors determining speech.

3.2 Psychological Considerations

Much of the research on the interview situation has been conducted by clinical psychologists, principally Matarazzo and Wiens (1972). After studying several groups of interviewees, and testing and retesting them in the interview environment, Matarazzo et al. (cited in Matarazzo 1973: 138) found very little intraspeaker variation. From this they concluded that the speech behaviour of any given individual is highly stable. They hypothesised that interviewer variables could influence the interviewee's average duration of utterance within certain parts of an interview (Matarazzo and Wiens 1972: 82). In fact, they found that an instructed interviewer can reproducibly modify, up or down and at will, the speech behaviour of one after another of his interviewees, and that as soon as he withdraws this influencing tactic the interviewees revert to their own baseline (Matarazzo and Wiens 1972: 118). If this is true for the sociolinguistic interview, then it has serious implications for the interviewer's technique on the language behaviour under study.

Other authors in the field of psychology report interesting results in the area of personality in the interview situation. For example, Cope (1969, cited in Furnham 1990: 80) found that Extroversion is the only trait which has consistently been found to be associated with a greater amount of verbal output or longer total speaking time.

4. Methodology

4.1 Participants[3]

Participants were chosen from a linguistics class at the University of Canterbury. Twenty-five individuals, including some non-New Zealanders, volunteered to participate. The relationship between the nationality of participants and their amount of talk time was considered, due to the possibility of different cultural conventions of a speech act (Gumperz 1982: 12). It was decided that nationality was not problematic in this case, as there were very few interviewees who were of non-New Zealand origin. It was also considered that nationality would not unduly affect the variable of personality, as there is evidence to suggest that the Big Five factors are reliable across cultures (Pervin 1993: 308–309).

Volunteers ranged in age from 19 to 68 years. This range was not considered problematic, as the Big Five have been found to be stable throughout life (McCrae and Costa 1990: 177). Indeed, Caspi and Moffitt (cited in McAdams 1994: 301) have found that 'trait consistency prevails even in the face of monumental changes in life circumstances.'

After consent was obtained, participants were requested to complete an NEO PI-R questionnaire form. The subjects were reminded that there are no right or wrong answers, and were requested to answer as honestly and accurately as possible. They were then interviewed by the author. Analysis of the NEO PI-R questionnaires was completed some time after the interviews, and the participants were given feedback in the form of a sheet summarising the results of their individual questionnaires.

A range of personality scores was obtained from the participants (see Table 1). In a normal population, the expected range for the majority is 30–70, the expected mean of T-scores is 50, and the expected standard deviation is 10. As can be seen from Table 1, this sample's means and standard deviations are close to the expected norms, therefore one can conclude that this sample does not differ markedly from the general population.

4.2 The Interview[4]

After completing the NEO PI-R form, each participant was interviewed individually in a room away from other participants. All interviews were conducted by the author and consisted of five questions:

- Have you ever been in a danger of death situation? Tell me about it.

	GREGARIOUSNESS	WARMTH	IDEAS	FANTASY
Mean	49.88	50.18	55.36	60.90
Median	54.08	51.05	55.60	61.60
Mode	54.26	51.05	61.60	67.60
Standard Deviation	12.66	11.98	9.31	9.93
Range	44.68	52.50	38.80	36.11
Minimum	24.47	21.75	33.60	41.49
Maximum	69.15	74.25	72.40	77.60
Sum	1246.91	1254.52	1384	1522.42
No. of participants	25	25	25	25

Table 1: Descriptive statistics of the facets of Gregariousness, Warmth, Ideas and Fantasy

- Have you ever had an experience involving UFO's, or aliens?
- Have you ever heard a good story about aliens/UFO's? Tell me about it.
- Have you ever seen a ghost?
- Do you know a good ghost story?

The interview was structured in the above way because I wanted to isolate the variable of personality in this research. I endeavoured to ask questions in the same way in every interview so that everything was kept constant apart from the variable of personality (see 'Discussion' section). It should be noted that in asking these five questions I wished to simulate the first five to ten minutes of a standard sociolinguistic interview so that I could look at personality in its 'rawest' form in the interview (i.e. before the participants 'warmed'[5] to the interview situation) and see why the beginnings of interviews and in fact, whole interviews are sometimes stilted. I acknowledge that this is the reason why many sociolinguists choose to ignore the first five to ten minutes of the interview for phonetic analysis purposes. However for this research it was appropriate to study the simulated beginnings of an interview to investigate personality, as the effect of personality sometimes 'wears off' during the course of a long interview. The point of looking at the interview in this way was to see if I could find out which personality traits are marked in speech.

The questions listed above are standard sociolinguistic interview questions

calculated to encourage respondents to become highly involved in *what* they are saying and thus pay less attention to *how* they are saying it. The first, Labov's famous 'Danger of Death' question, is one frequently used in sociolinguistic research to overcome the 'Observer's Paradox' and gain access to the vernacular (1978: 209-210).

The interviews typically lasted between three and ten minutes, and the total duration of the interviews ranged from one minute to twenty-three minutes.

4. Analysis

Three types of analyses were employed to define the notion of 'quantity of speech'. The first focused on the participants' percentage of talk time versus their Warmth, Gregariousness, Ideas, and Fantasy scores. The second analysis considered the interviewer's speech and pause time versus the interviewee's speech and pause time. The third analysis involved an examination of the way in which the interviewees said 'No' to the set questions.

Speech and pausing was timed using a stopwatch. Pauses as well as speech were included when calculating the interviewee speech duration, as these two measures combined yield a 'speech turn'. The total duration of the interview was also measured, and the proportion of 'talk time' for the interviewee, including pauses, was calculated as a proportion of the total interview time.

Normality of the data was checked using the Wilk-Shapiro test, which tests for normality when the number of subjects is less than fifty. After determining that the data was not normally distributed, the non-parametric test Kendall's tau-b was used.

In the qualitative analysis, coded descriptions were assigned to the ways in which the interviewees said 'No' to the set interview questions.[6] The codes used in this study are ordered in terms of length below.

SHORT = 'No', 'No' response, with long pauses, 'No. Not really' answer, and 'No', then curt explanation.

MEDIUM = 'No', then short explanation, 'No', then explanation.

LONG = 'No', then long explanation.

5. Results

In addition to Kendall's tau-b analysis, Bonferroni correction was used with the data, because testing four different facets draws on the dataset four times, and this must be accounted for when investigating significance levels. After Bonferroni correction, for the results investigating facets to be significant at an overall level of $p < 0.05$, p should be less than 0.0125. Table 2 gives the Kendall's tau-b results for all of the relevant comparisons.

Table 2 shows that, with regard to the amount of talk, there were only two statistically significant results in this study. Firstly, despite my best efforts to keep it uniform across the interviews, the duration of my speech actually varied, and the results show a significant relationship between interviewer's talk time and interviewee's talk time (tau-b = 0.480, p =0.001).

Secondly, there is a strong relationship between participants' percentage of talk time and Warmth scores (tau-b = 0.363, p =0.013). Warmth is the only one of the Extroversion facets that is significant. The Openness facets investigated are not significant in predicting quantity of interviewee speech, in spite of the type of question asked in the interview, which related to those specific facets.

With regard to the analysis of how participants said 'No', the results show

COMPARISON	KENDALL'S TAU-B	P =
Participants percentage of talk time versus Warmth score	0.363	0.013
Participants percentage of talk time versus Gregariousness score	0.143	n.s
Participants percentage of talk time versus Ideas score	-0.017	n.s
Participants percentage of talk time versus Fantasy score	0.068	n.s
Interviewee duration versus interviewer duration	0.480	0.001

Table 2: Kendall's tau-b analysis of talk time versus other factors

SUBJECT	W SCORE	G SCORE	'NO' RESPONSE TYPE	UNDER/OVER MEAN FOR WARMTH	UNDER/OVER MEAN FOR GREGARIOUSNESS
U	21.75	37.76	SHORT	UNDER	UNDER
C	32.63	24.47	SHORT	UNDER	UNDER
M	32.63	35.11	SHORT	UNDER	UNDER
D	32.63	43.62	SHORT	UNDER	UNDER
W	37.89	47.87	SHORT	UNDER	UNDER
V	39.25	35.71	SHORT	UNDER	UNDER
Q	49.25	43.88	SHORT	UNDER	UNDER
T	48.42	60.64	MEDIUM	UNDER	OVER
E	51.05	43.62	MEDIUM	OVER	UNDER
J	51.05	54.26	MEDIUM	OVER	OVER
F	51.05	54.26	MEDIUM	OVER	OVER
P	51.05	67.02	MEDIUM	OVER	OVER
L	53.68	35.11	MEDIUM	OVER	UNDER
B	53.68	43.62	MEDIUM	OVER	UNDER
A	54.25	52.04	MEDIUM	OVER	OVER
X	56.32	54.26	MEDIUM	OVER	OVER
S	48.42	56.38	LONG	UNDER	OVER
O	48.42	56.38	LONG	UNDER	OVER
I	56.32	67.02	LONG	OVER	OVER
N	58.95	69.15	LONG	OVER	OVER
H	61.58	62.77	LONG	OVER	OVER
K	64.21	67.02	LONG	OVER	OVER
R	74.25	54.08	LONG	OVER	OVER

Key: W= Warmth G= Gregariousness
Mean Warmth Score = 50.18 Mean Gregariousness Score= 49.88
N.B: Two of the participants interviewed for this research replied "yes" to every question and are therefore not included in this table.

Table 3: Interviewee's types of 'No' response (in order of response length)

a clear pattern (See Table 3). Interviewees whose Warmth and Gregariousness scores were below the mean gave minimal (short) answers to the interview questions when replying in the negative. Minimal answers are defined as 'No', '"No" with long pauses', 'No, not really' and '"No" plus curt explanation'. Speakers with either a Warmth or a Gregariousness score above the mean regularly gave longer negative responses, such as '"No", then short explanation', '"No" then explanation'.

For definitions of 'No' response coding, please refer to Analysis section above.

6. Discussion

The results of this research indicate that the way personality affects the sociolinguistic interview situation is complex. It was expected that facets of Extroversion, namely Warmth and Gregariousness, would be strongly related to quantity of interviewee speech, but in fact only Warmth was seen to be an important factor in this respect. It was also expected that the nature of participants' 'No' responses would be related to their Warmth and Gregariousness scores. Indeed, those participants whose scores on these facets exceeded the mean regularly gave longer responses compared with those participants with Warmth and Gregariousness scores below the mean. Therefore, although statistically insignificant in relation to interviewee duration, Gregariousness does appear to interact with the facet of Warmth to affect the nature of participants' interview behaviour.

Results on the Ideas and Fantasy facets of the Openness domain were found to be non-significant, indicating that this domain is less related to quantity of speech than Extroversion. These facets were investigated as the interview questions were indirectly involved in the 'Openness to Experience' facets, Ideas and Fantasy. It was thought that a person with a low score in 'Fantasy' or 'Ideas' would probably talk for less periods of time on this subject, as they were not open to the paranormal. In fact, it was found that there was no relationship between 'Openness to Experience' facets and amount of speech from the interviewee, therefore showing that the questions employed in this study had no effect on the response given by the interviewee.

With regard to the significant relationship found between the duration of interviewer speech and the duration of interviewee speech, the initial hypothesis in this study was that it was unlikely that a strong association would be

found. It could be argued that this result does not demonstrate causation in one direction or the other; it merely shows that there is an association. Matarazzo and Wiens (1972: 118), in their interviewee modification study, found that the interviewer's talk time can influence the interviewee's talk time. These researchers were, of course, deliberately modifying the interviewer utterance to test their hypothesis that the interviewee is influenced by the interviewer. However, in the present study the same interviewer (the author) conducted all the interviews and actively attempted to keep speech and pause time consistent across interviews. In fact, the interviewers talk time varied and it can be concluded that the interviewer was influenced by the interviewees, and not the other way around as Matarazzo and Wiens (1972) have found.

Such a significant relationship indicates an interaction of the interviewer's and interviewee's personality traits, something which should be investigated further in the future. The result demonstrates clearly the extent to which sociolinguistic interviewers can be subconsciously influenced by the person they are interviewing. Like other individuals, sociolinguists exhibit the phenomenon of 'convergence' which occurs in a social interaction where individuals shift their speech styles to become more like that of those with whom they are interacting (Giles and Smith 1979: 46). Giles and Coupland (1991: 63) define convergence as 'a strategy whereby individuals adapt to each other's communicative behaviours in terms of a wide range of linguistic/prosodic/non-vocal features including speech rate, pausal phenomena and utterance length, phonological variants, smiling, gaze and so on.'

This is important for the linguistic interview situation, as convergence has implications for the amount of speech obtained for the purposes of linguistic analysis. For example, in the situation where a talkative and non-talkative participant are interacting, one might expect that the talkative person may reduce their utterance length to become more like that of the non-talkative person. This is, in fact, what happened in this research. My own speech and pause time (interviewer duration) was significantly influenced by the interviewee's speech and pause time (interviewee duration). Personality therefore, can account for why the interviewee in a given interview does not speak much. I believe that it is not the case that this lack of speech directly reflects the interviewer's ability in that situation, it is merely that personality is a stronger factor. Another point which is worth mentioning is that perhaps the style of the interviewee's speech relates to personality and therefore if a casual/formal style is required in a given study it may be wise to include those speakers who have high and low scores in 'Warmth'.

Further study to investigate the phenomenon of personality interaction between both the interviewee and the interviewer would be worthwhile, as this may provide more clues as to how personality can be an influential factor in the sociolinguistic interview situation with respect to quantity of speech (and indeed other areas such as phonology, style, etc). Future studies may also find that using different questions, approaching the interview differently, or looking at different facets of personality as identified by the NEO-PI, gains different results to those discussed above, as it is possible that other personality facets could have an influence on the interview situation.

This study has been worthwhile as we can now see that it is indeed possible to study personality in the sociolinguistic interview situation. This study has also been important as it provides more understanding of what is involved in the sociolinguistic interview situation. I believe that there is still much to be researched on this matter, and that with more research, we will gain better understanding of personality in the sociolinguistic interview situation instead of merely taking it for granted that personality affects the sociolinguistic interview. As we now know, personality is a factor in this type of situation, and, more specifically, we know that 'Warmth' affects this situation. Other traits may be investigated in the future to determine specifically which traits affect the sociolinguistic interview.

7. Conclusion

Social scientists have speculated about the importance of Extroversion in a two-person interaction, but until now research has not been conducted on specific facets. The results of this study show that the interviewee's personality—in particular the Warmth facet of Extroversion—influences the amount of talk. A further finding is that the nature of participants' 'No' responses relates to their Gregariousness and Warmth scores.

One outcome of this study that was not entirely expected was the significant relationship between interviewer duration and interviewee duration, this despite the fact that the interviewer was trying to be consistent with all participants. This shows that 'convergence' was occurring in the interviews, and perhaps even 'convergence' of the interviewer's and the interviewees' personalities.

This research was carried out to see how personality affects the sociolinguistic interview situation. More research on this area will provide linguists and psychologists with sound evidence on which traits influence speech quantity

in a given interaction. The sociolinguistic interview situation is complicated, however I have demonstrated that the effects of personality on the interview can be tested.

Notes

I would like to thank Margaret Maclagan of the Speech and Language Therapy Department at the University of Canterbury, for supervising this research project, and also for her helpful comments on earlier versions of this paper. I would like to acknowledge the support of Gary Steel, Department of Human and Leisure Sciences at Lincoln University, for his advice on personality measures and statistical analyses. I would like to thank Gillian Lewis for helping to edit this paper. Finally, I would like to thank the participants of this research, as without their involvement this study would not have been possible.

1 The phenomenon of the sociolinguistic interview situation will not be discussed here. Discussions of this type of interview have been made by many linguists and social scientists (See, for example, Milroy 1987; Wolfson 1976; Labov 1978; Schiffrin in Tannen 1993; Gumperz and Hymes 1972).
2 See Bell and Johnson (1997) for a study on gender and ethnicity, and Cukor-Avila and Bailey (2001) for a study on ethnicity in the sociolinguistic interview situation.
3 Approval for this research was given by the Human Ethics Committee at the University of Canterbury. Great care was taken to assure participants of their anonymity, and to ensure that all ethical considerations were considered and resolved.
4 Participants in this research were given a code number for the interview, and a different code number for the NEO-PI personality test to guarantee their anonymity. As a further precaution, they were subsequently assigned a different code again, for the write-up of this research.
5 'warmed' is used here in the general sense.
6 If participants said 'yes' to the questions in the interview, they always provided an explanation. For this reason, analysis was carried out to investigate if there were patterns of how the interviewees said 'No'. Any participant who consistently answered positively to the interview questions was not included in the 'No' analysis, for obvious reasons. The results were checked independently and both the author and the independent checker agreed on the results as shown in Table 3.

References

Allport, Gordon W. and H.S. Odbert. 1936. 'Trait names: A psycho-lexical study.' *Psychological Monographs* 47 (Whole No. 211).

Bell, Allan and Garry Johnson. 1997. 'Towards a sociolinguistics of style'. University of Pennsylvania Working Papers in Linguistics *4 (1): 1-21*.

Brown, Penelope and C. Fraser. 1979. 'Speech as a marker of situation.' In Klaus R. Scherer and Howard Giles (eds) *Social Markers in Speech*. Cambridge: Cambridge University Press.

Cattell, Ray B. 1946. *The Description and Measurement of Personality*. New York: World Book.

Costa, P.T. Jr. and R.R. McCrae. 1992. NEO PI-R: *Professional Manual*. Odessa: Psychological Assessment Resources.

Cukor-Avila, P. and G. Bailey. 2001. 'The effects of the race of the interviewer on sociolinguistic fieldwork'. *Journal of Sociolinguistics* 5 (2): 254-270.

Furnham, A. 1990. 'Language and Personality.' In Howard Giles and Robinson, W.P. (eds) *Handbook of Language and Social Psychology*. Chichester: Wiley.

Giles, Howard and Nikolas Coupland. 1991. *Language: Contexts and Consequences*. Milton Keynes: Open University Press.

Giles, Howard and W.P. Robinson. 1990 *Handbook of Language and Social Psychology*. Chichester: John Wiley.

Giles, Howard and P. Smith. 1979. 'Accommodation theory: Optimal levels of convergence.' In Giles, Howard and R. N. St Clair. (eds) *Language and Social Psychology*. Oxford: Basil Blackwell.

Gumperz, John J. 1982. *Language and Social Identity*. Cambridge: Cambridge University Press.

Gumperz, John J. and Dell Hymes. 1972. (eds) *Directions in Sociolinguistics: The Ethnography of Communication*. New York: Holt, Rinehart and Winston.

Labov, William. 1978. *Sociolinguistic Patterns*. Oxford: Basil Blackwell.

Matarazzo, J.D. 1973. 'A speech interaction system.' In D. J. Kiesler. (ed) *The Process of Psychotherapy: Empirical Foundations and Systems of Analysis*. Chicago: Aldine Publishing Company.

Matarazzo, J.D. and A.N. Wiens. 1972. *The Interview: Research on Its Anatomy and Structure*. Chicago: Aldine Atherton.

McAdams, D.P. 1994. 'Can personality change? Levels of stability and growth in personality across the life span.' In T. F. Heatherton and J. L. Weinberger (eds) *Can Personality Change?* Washington D.C: American Psychological Association.

McCrae, R.R. and P.T. Costa Jr. 1990. *Personality in Adulthood*. New York: The Guilford Press.

Milroy, Lesley. 1987. *Observing and Analysing Natural Language*. Oxford: Basil Blackwell.

Pervin, L. A. 1993. *Personality: Theory and Research* (6th ed.). New York: Wiley.

Rickford, John R. and Fay McNair-Knox. 1994. 'Addressee- and topic-influenced style shift: A quantitative sociolinguistic study.' In Douglas Biber and Edward Finegan. (eds) *Sociolinguistic Perspectives on Register*. New York: Oxford University Press.

Robinson, W. P. 1972. *Language and Social Behaviour*. Middlesex: Penguin Books.

Scherer, Klaus R. 1979. 'Voice and speech correlates of perceived social influence in simulated juries." In Howard Giles and R.N. St Clair. (eds) *Language and Social Psychology*. Oxford: Basil Blackwell.

Scherer, Klaus R. and Howard Giles 1979. *Social Markers in speech*. Cambridge: Cambridge University Press.

Tannen, Deborah. 1993. *Framing in Discourse*. New York: Oxford University Press.

Wolfram, Walt and Ralph Fasold. 1974. *The Study of Social Dialects in American English*. New Jersey: Prentice-Hall.

Wolfson, Nessa. 1976. 'Speech events and natural speech: Some implications for sociolinguistic methodology.' *Language in Society* 5: 189-209.