TE REO

Journal of the Linguistic Society of New Zealand

Volume 58 *2015*

Contents

ACQUIRING SOME *LIKE*-NESS TO OTHERS: HOW SOME POLISH TEENAGERS ACQUIRE THE SCOTTISH PRAGMATICS OF *LIKE*

Sarah Truesdale: *School of Linguistics and Applied Language Studies, Victoria University of Wellington <sarah.truesdale@gmail.com>*

Miriam Meyerhoff: *School of Linguistics and Applied Language Studies, Victoria University of Wellington <miriam.meyerhoff@vuw.ac.nz>*

Abstract

Numerous studies have examined uses of *like* as a discourse marker (DM) in vernacular English. However, we know little about how L2 English speakers learn the norms associated with it. The functions of *like* are not explicitly taught and must be picked up through exposure. The distribution of *like* in L2 speech provides a unique window into how learners can or cannot successfully match patterns in the speech around them. We examine the acquisition of *like* among Polish teenagers who have moved to Scotland, comparing the frequency and functions of *like* in their speech with the speech of same-age Scottish L1 speakers. L2 speakers take some time to master the range of pragmatic functions of *like*, with the most common functions acquired first; the distribution of *like* is constrained by only linguistic factors in L1 speakers, but is constrained by both social and linguistic factors in the learners.

1. Introduction

This paper is a study of language variation and of language acquisition.[1] It draws on a corpus of spontaneous speech produced by Edinburgh adolescents to investigate variation in the use of the discourse marker (DM) *like* by non-native speakers of English. We compare the way Polish-born teenagers living in Scotland use *like* in their conversation with the way their locally-born classmates use *like*. The uses we are interested in range from filler tokens of *like,* to hedging uses, through to uses of *like* as a quotative verb.

Following Meyerhoff (2009), we consider three possible outcomes for the transfer of variation in the use of DMs in the speech of one group of speakers to the speech of another group of speakers:

1. No transfer (failure to acquire anything like the L1 patterns of variation);

2. Transfer of the L1 form but failure to acquire the L1 constraints on the variation;

3. Transfer of both the L1 form and the L1 constraints.

The field of studying variation in DMs has, as yet, made only limited forays into the methods and approaches of variationist sociolinguistics (Sankoff et al. 1997 is a notable and early exception; Cheshire 2005, 2007 follow her earlier work extending variationist methods into new domains of language use and there has recently been a flurry of interest in *like* – more on which shortly). Moreover, as far as we are aware, there has been next to no attempt to engage this study with the field of contact linguistics (though Matras 2012 certainly discusses DMs in contexts of language contact). The data presented in this paper seeks to move forward in both these directions, outlining methods and issues that researchers may need to take into consideration when looking at discourse markers. We will differentiate between form, frequency, and relative frequency of form across different functions. Specifically, we ask:

1. Do the Polish-born teenagers in the corpus use the DM *like*?

2. If they do,
    a. do they use it with the same frequency as their locally-born peers?
    b. do they use it for the same discourse functions?

  c. is the variation in their use of *like* subject to the same social and linguistic constraints as the variation in locally-born speakers' use of *like*?

The paper is structured as follows. First, we provide a little background motivating the treatment of *like* as a discourse marker, explaining how we are using the term. Second, we provide some background on our data and methods. We introduce the corpus in more detail and explain what a variationist analysis of *like* entails and what kinds of questions it is suited to posing and answering. Third, we present our results and analysis of the results.

2. *Like* as a discourse marker

In this section, we examine the notion of *discourse marker* and motivate the analysis and treatment of *like* as a DM as a prelude to introducing the corpus and methods we have used to study the use of *like* in two groups of speakers.

 There does not seem to be a general consensus as to what exactly constitutes a discourse marker. They were once considered to be a 'peripheral phenomenon' (Bazzanella et al. 2007: 10), not truly part of the grammar, but this view has been challenged by successive generations of pragmatics researchers. DMs are now usually treated as a set of items belonging to one heterogeneous functional category. However, this heterogeneity poses some special problems, not least of which is how to define them in a manner that allows us to reliably identify a form as a member of the class of DMs. Although 'there is considerable variation in what might be labelled Discourse Markers' (Fraser 2009: 294), the following five characteristics are commonly agreed upon defining features. (Since the study of discourse markers and variation has, arguably, been characterised by a lack of historicity, we focus on the definition of DMs in such a way that highlights the different components' pedigree. Pichler 2010 provides an alternate historical perspective on the emergence of consensus on how to deal with variation in DMs.)

1. *DMs are markers which manage discourse. They do this by signalling*
a relationship between the different parts of the discourse.
Fraser (1990), for instance, says that DMs are expressions 'which signal a sequential relationship between the current basic message and the previous discourse' (1990: 383). Similarly, Hansen (1995) defines discourse markers

as sequentially dependent, pragmatic markers which indicate a relationship between what follows and the previous discourse (cf. Schiffrin 1987). Hansen argues that they are a kind of connective 'whose scope is in principle free, in so far as they may connect units below and above the level of the utterance and may even connect to extralinguistic elements' (Hansen 1995: 32). Agreement on this fundamental property has remained stable across time, for example, Hopper and Traugott also define discourse markers as markers that 'serve to manage the segments of the discourse' (2009: 129).

2. *They are subject to syntactic constraints but do not create, nor are they subject to, syntactic dependencies.*
This characteristic expresses a crucial property of DMs, namely, that DMs can be omitted without making a sentence ungrammatical (Hansen 1995: 32), cf. Schiffrin's (1987) notion that they are sequentially dependent and bracket other talk. DMs are, of course, not the only clausal constituents this is true of (adverbials are also generally free of dependencies – a few verbs such as *put* require an adverbial argument but this does not negate the point as a whole), but alongside the other criteria, this one can be helpful.

3. *They do not affect the propositional content of an utterance.*
DMs 'do not refer, nor are they capable of affecting truth-conditions' (Hansen 1995: 32, cf. Blakemore 2002). This criterion merits a little closer scrutiny. We understand Hansen to mean that (a) a DM has no referential content in and of itself, and (b) if a speaker uses a discourse marker like *I mean* or *kinda*, this does not change the conditions under which the proposition or constituent under the scope of the DM is true or false. That is, the smaller linguistic units, such as NPs and VPs, as well as larger units such as clauses that refer to something (or in any way predicate *t* of an *e*) continue to so refer even when under the scope of a DM.

However, some semanticists have argued that hedges (at least, among the larger class of DMs) can be represented as interacting with the propositional content of an utterance. Rosch's (1973, Rosch and Mervis 1975) influential work on cognitive prototypes showed us that it is valid to say some colours and shapes are 'better' or 'worse' examples of the basic colour or shape term than others. Lakoff's (1973) exploration of this showed that the effect of hedges like *sort of, pretty* and *technically* could be argued to affect the truth conditions of an utterance, i.e. if John is 185cm, the authors would want to say that it is true that *John is tall*; but if John is 185cm, we would want to say

that *John is sort of tall* is false (from our perspective, he isn't 'sort of tall', he 'really **is** tall'). Given this, and building on fuzzy set theory (Zadeh 1965), Lakoff proposed a fuzzy logic which would allow us to specify the interaction between the value of the hedge, the proposition and the felicity conditions for uttering the proposition and general rules of conversation (1973: 490).

Similarly, DMs that scope over whole propositions such as *of course, I mean* do not change the truth theoretic status of the proposition(s) they scope over. Nevertheless, it has been claimed that some such DMs introduce information that does involve an interaction between propositional semantics and assertions about the respective stance of the speaker and addressee or they predicate over possible worlds. Cresswell (1990) and Clift (2001) arrive at similar conclusions using very different methodologies. (Incidentally, it is not at all clear how this might be represented within a Neodavidsonian semantics (Kearns 2000 §8.2), perhaps to some readers this will indicate a limitation of such perspectives on semantics, but Rini and Cresswell (2012) provide one way of representing the issues with modality that some such discourse markers raise.)

Both features 2 and 3 mean DMs are problematic as the object of study from a variationist perspective for reasons we will outline shortly.

4. *The function of DMs is instructional or operational.*
In using a DM, the speaker gives instructions to the hearer 'on how to manipulate the conceptual or propositional content of the stretch of discourse marked, with a view to integrating it into a coherent mental model' (Hansen 1995: 32).

This property is a corollary of 3. The contribution to meaning that a DM makes is generally at an interpersonal and interactional level. Another way of expressing Hansen's point would be to say that DMs provide information about the alignment of the interlocutors' common ground. Some DMs are specialised to provide information about the non-alignment of the interlocutors cf. discussions of *oh*-prefacing (Heritage 1998). Heritage (2013) provides an excellent overview of this literature cross-linguistically.

5. *DMs are polyfunctional.*
Hansen (1995, 1998, 2005) notes that discourse markers tend to be multifunctional. For example, looking at French DMs, Hansen identified 'no less than 14 different uses' (2005: 154) of the adverb *enfin* 'finally', when it acts as a DM.

Schourup sums up a typical DM by saying that it is 'a syntactically optional expression that does not affect the truth-conditions associated with an utterance it introduces and is used to relate this utterance to the immediately preceding utterance' (1999: 234). We endorse Schourup's definition with a slight amendment allowing for DMs to relate utterances beyond the immediately preceding linguistic context.

2.1 Discourse functions of like

Like as a discourse marker is arguably 'one of the most salient features of present-day vernacular English' (D'Arcy 2005: 2). It attracts interest for many reasons, including its transnational spread and its 'deep roots in the history of English' (D'Arcy 2005: 5). A number of linguists have, over the years, turned their attention to *like* and attempted to describe its possible functions.

The use of *like* as a DM is not novel: Miller and Weinert note that the Oxford English Dictionary has entries for 'dialectal and vulgar' (1995: 367) uses of *like* from as far back as the early nineteenth century where it was 'used parenthetically to qualify a preceding statement: = "as it were", "so to speak"' (1995: 367). Until very recently, they observe, *like* has been described as a 'meaningless interjection or expletive' (1995: 367), a feature that is essentially redundant (recall the characteristics of DMs introduced in the last section), vague, a hedge, or a tool which engenders a toning down effect (1995: 368). However, Miller and Weinert state that many of the authors who used these descriptions do not provide a rich textual context for the uses described and such context is essential for evaluating the interactional functions that *like* as a DM may serve.

Since then, various uses of *like* have been documented. Schourup (1985) notes that it is particularly common as a focuser before numerals, and Underhill (1988) observes that *like* as a focus particle occurs in a range of syntactic positions, and the main constraint on its distribution is that it must accompany new information. Brinton (1996) takes issue with this pragmatic constraint, pointing out that not all uses of *like* are associated with new information. It seems that in some cases, the use of *like* may be more of an index that the information it modifies is somehow novel, unexpected or indicates a new (re-)alignment of the interlocutors' attention/common ground (hence the use of *like* with self-repair). Levey (2003: 29) describes it as serving 'affective and solidarity-building purposes'. Most of the prior work has drawn on data from North American English though Miller and Weinert (1995) summarise the distribution of *like* in British English, including the distinctively British

clause-final use of *like*. They document *like* as an approximation before numerals, before direct discourse (introducing direct speech), as a clarification after a question, as a marker when giving examples and as an interjection (Miller and Weinert 1995: 369).

Clearly, given these observations about the use of *like*, it satisfies several of the definitional criteria for a discourse marker, as outlined above. Specifically, it is polyfunctional; it is not a syntactic dependency; it is a marker of new or changed information states; it does not contribute to the truth-conditional semantics of the utterance, instead relating parts of the discourse to each other, or expressing 'operational' information.

Why the quotative function of *like* is to be studied alongside the focus functions remains to be motivated, since it is less clear how well quotative *like* satisfies our definitional criteria for DMs. We decided to group it together with the other functions of *like* because it seems fairly clear that the use of *like* as a quotative verb was made possible by the similative or approximative functions of *like* elsewhere in the grammar. Buchstaller (2014) traces the emergence of the English quotatives, including *(be) like* and shows that there are parallel developments of approximative and similiative elements in many languages world-wide with such typological, temporal and geographical spread that it would be perverse to assume they are all calques of the English quotative use of *like* (cf. Table 1.2, p.20). Since this paper is interested not only in documenting the constraints on the use of *like* in the speech of the locally-born Edinburgh teenagers, but also in documenting the extent to which L2 teenagers from Poland are acquiring the same kinds of norms, we would like to know whether the acquisition patterns do or do not mirror the historical development of the functions of *like* in English. (This is familiar question of whether ontogeny recapitulates phylogeny.) In other words, the inclusion of *like* as a quotative enables us to consider the extent to which the enrichment of the variable in the speech of L2 users is the same or different from the enrichment of the functions of *like* over time.

2.2 Discourse markers as variables
Many of the properties of DMs make them potentially problematic variables. Clearly, their optional nature means that they deviate from the binary presence/absence distinction of the classic sociolinguistic variable. In addition, the early analyses of variation considered phonic variables only (Labov 1972, Cedergren and Sankoff 1974; we follow Walker 2010 for the terminology) and a basic tenet of the variationist method was that the alternation between

variants was semantically vacuous. For example, there is no semantic load associated with different pronunciations of a word: a speaker's use of [dɑt] or [dɑɹt] for *dart* has no denotational impact. The lack of any semantic load was initially treated as criterial, along with strict structural occurrence restrictions, so the presence or absence of a constricted /r/ in *dart* only applies to non-prevocalic tokens of /r/: one cannot alternate [dɑt] and [dɹɑt]. Yet, as we have seen, DMs are both optional and carry some meaning (albeit pragmatic rather than semantic).

However, the methods of comparing equivalent variants quickly expanded and syntactic variables also became the focus of variationist study from the 1970s on (Wouk 1999, Cheshire 2005, 2007 helpfully rehearse the main issues, Pichler 2010 also). Some such variables included the alternation between active and passive in English, the use of subjunctive versus indicative in Spanish and the alternation between être and *avoir* auxiliaries in French (Sankoff and Thibault 1977). The notion of a variable expanded, in other words, to encompass forms that are substitutable even if there is some meaningful difference between the variants chosen (as with the active and passive, which give different prominence to the agent of an event). Notice though that even if there is some kind of semantic load associated with the different variants, the structural restriction remains intact: a speaker may sometimes say *il est monté* ('he climbed') and sometimes *il a monté*, but the choice of auxiliary must fill the same slot; the choice of indicative versus subjective inflections on a Spanish verb always occur in the same position immediately following the verb stem; and the alternation between active and passive in English is subject to very strict and predictable formal constraints on argument position and verb form. Walker (2010) provides a concise account of how the notion of 'variants' of a single variable has been extended to allow for functional equivalence, opening the door not only for syntactic variables such as the ones mentioned here.

However, DMs often display a startling degree of structural promiscuity. It sometimes appears that they can occur freely virtually anywhere in the clause (consider the potential distribution of downtoner hedges like *sort of*, Holmes 1988, the particle *eh* in New Zealand English, Meyerhoff 1994). Indeed, this is a general characteristic of features that express politeness or intersubjectivity. This structural promiscuity poses a variationist problem of a somewhat different order to the equivalence problem that had to be resolved for grammatical variables. This is because there is no tidy envelope of variation for DMs that the researcher can define.

Nevertheless, a number of linguists have shared the intuition that the presence/absence of discourse particles may also be socially and linguistically constrained, just as phonic or grammatical variables are. The problem with not having a clear envelope of variation is a quantitative one: what is the denominator going to be when you can't specify all and every place a speaker might choose to express politeness or intersubjective alignment? For example, expected frequencies of a hedge balloon to ridiculous proportions if we were to try and quantify the occurrence of *sort of* as potentially occurring before any XP (which seems to be its distributional potential in at least some varieties of English).

Moreover, even if we were to do this, it is debatable whether it is warranted. Although the distribution of *like* can in theory occur before many different XPs (but perhaps not before personal pronouns or NPs in idioms (Wohlgemuth 2003, cited in D'Arcy 2005)), there are obvious constraints on co-occurrence. Consider the sentence *I heard their train is running 15 minutes late.* In theory, *like* can precede every lexeme in this sentence (and in some dialects can also occur clause-finally). Although some of these slots sound more natural when filled than others do (and this is of course one of the facts we would like to explain), a speaker would sound nothing short of pathological if they were to use *like* in every possible slot: **Like I like heard like their like train like is like running like 15 like minutes like late, like.* This means that a model of the variation which starts from the presumption that every slot is potentially fillable is of dubious value.

There are different ways of dealing with this within a variationist paradigm. D'Arcy's (2005) analysis of *like* in Canadian English resolves this with sampling: where the envelope of variation is so large as to be difficult to operationalise, she sampled selectively. She extracted 75 randomly selected CPs per speaker for an analysis of clause-initial *like*.[2] In effect, the position adopted in this approach is to control the denominator in a way that analyses of variation taken from naturally occurring conversations usually cannot, thereby allowing for comparability across speakers. The cost is at the expense of full accountability to the numerator (cf. Schegloff 1993).

Early research on politeness phenomena (Preisler 1986: 58, Schiffrin 1987: 8, 13, Schegloff 1988, Tannen 1988, Holmes 1988, 1989: 297) resolved the problem somewhat differently, sometimes by adopting qualitative methods of analysis. Another solution, when it may not be possible or practical to specify the denominator accurately or exhaustively, is to create some other form of index that is comparable across speakers. Meyerhoff (1994) quantified the

number of tokens of a pragmatic particle over the number of words produced, creating a frequency index. This solution enables us to say something about high and low frequency users (out of the total sample) of a particular DM and high and low frequency functions (relative to all functions of the DM), but since it is at the expense of linguistic detail, we cannot say much about how a DM interacts with the syntax of the rest of the clause. It therefore works well if the primary questions are about the social and transactional meanings of a DM. Indeed, this is the approach adopted by Nestor (2013).

In this study, we are focusing on the functions of *like*. This enables us to undertake a quantitative analysis without being overwhelmed by the volume of data. Two considerations informed this decision. First, we believe that the interactional nature of a hedge like *like,* its polyfunctionality and its ambiguity with respect to its own grammatical category mean that learners acquiring the form naturalistically through conversation and ambient exposure are likely to be oriented initially to functions. Second, some of the Polish teenagers are very low users of *like* and sometimes rather basic users of English. By focusing on a functional analysis, we felt we were more likely to be able to perceive emergent similarities between the Edinburgh- and Poland-born teenagers. Whether functions of DMs are learnt before the syntax associated with them is an empirical question and we look forward to further research that will test this presumption.

3. Data and methods

Our data is taken from conversational interviews conducted and recorded for the *Sociolinguistics and immigration: linguistic variation among adolescents in London and Edinburgh project* (ESRC RES 000-22-3244). The sample used for this study consisted of utterances from teenagers aged between fourteen and seventeen years. Tokens were extracted from a convenience sample of the conversations of five Scottish girls and five Scottish boys (from a total of thirteen girls and eight boys) and five Polish girls and five Polish boys (from a total of eight girls and eight boys). Our Polish teenagers were from varying socioeconomic backgrounds based on their parents' occupations. All had moved to Edinburgh with their families, wanting to take advantage of the economic opportunities that came with the opening up of the UK labour market following the expansion of the EU in 2004. The Polish adolescents had been living in Scotland for between nine months and four years and they

reported that they had been learning English for between nine months and seven years.

It is possible that this masks even more variance than it would first seem. Anecdotally, we understand that the quality of English teaching varied considerably in different parts of Poland in the first decade of the century. Participants were mostly from urban backgrounds where we are told English teachers were more likely to be trained in English language teaching, but in rural areas, we are told this was not necessarily the case (Agata Daleszyńska, Anna Strycharz, Emilia Wrọbel pers. comm.).

At any rate, we consider the formal language instruction that the Polish teenagers had to be less relevant than their experience in Edinburgh, because the DM uses of *like* that we are interested in are not explicitly taught and have to be acquired through more implicit cues and naturalistic learning.

Our investigation proceeded with the hypothesis that the Polish adolescents will show evidence of adopting *like* in order to integrate locally, but the range of functions with which they use *like* will be more limited than native speakers, depending on:

☐ Friendship network (mainly Polish versus mainly Scottish or mixed)

☐ Length of time spent learning English

☐ Length of time spent in the UK

Following Levey (2003), it is also hypothesised that there will be gender differences in the use of *like*, that is, female students will use *like* with greater frequency and across a wider range of functions than males 'for [the] affective and solidarity-building purposes' (2003: 29).

3.1 Data handling
Tokens of the discourse marker *like* were taken from conversations from the 'Sociolinguistics and Immigration' corpus. As mentioned earlier, this data was narrowed down to ten Polish adolescents and ten Scottish adolescents, with an even split of genders.

The data was explored and analysed using the multiple regression functions available in Rbrul (Johnson 2009). It is beyond the scope of this paper to explain multivariate regression analysis thoroughly. Walker (2010) is a remarkably clear introduction to the conceptual and practical issues associated with using multiple regression to test linguistic hypotheses. Johnson (2009)

and Paolillo (2013) are more detailed discussions and address the advantages and disadvantages of the two programmes. Essentially, the purpose of using multivariate analysis is to allow us to consider the relative effect of different possible predictor factors on a dependent variable. This is particularly helpful when our data is distributed unevenly across speakers and across contexts within the speech of one person, as DMs almost always are.

Generally, the approach requires us to test the application or non-application of a hypothesised (variable) rule in order to determine what contexts more or less strongly favour or disfavour the occurrence of the variant we are interested in. In the analysis that follows, we use the methods to ask questions that will enable us to compare the way in which the Edinburgh-born and the Poland-born teenagers are using *like* as a DM in the recorded conversations. We use multiple regression as a tool for exploring the importance of different linguistic and non-linguistic factors in determining how the teenagers in our corpus use *like*.

3.2 The dependent variable: like

Only tokens of *like* as a DM were counted. Tokens of *like* functioning as a verb (e.g. *I like chocolate cake*), as a preposition (e.g. *He has a car like mine*) or as a suffix (e.g. *She is very childlike*) were not counted.

Practice differs in the handling of DMs for variationist analysis. Some research focuses principally on formally interchangeable structures with similar semantics (cf. Dines' 1980 classic work on general extenders); some focuses on interchangeable functions. Pichler (2010) concludes, in line with existing practice, that different methods may be appropriate for different DMs. For this study, we have focused on the form *like* since (as the research questions stated) our interest is in documenting the extent of parallelism between the L1 and the L2 speakers' use of the lexeme. Several different functions of *like* were coded for: focus; word-finding difficulties; quotative; approximation (but see below); clause-final modifier; corrective and unclear/other.

Given that discourse markers can have a variety of functions and can even 'exercise two or more functions simultaneously in a given piece of text' (Miller and Weinert 1995: 372), it was not always easy to identify the function of any given token of *like*. We have opted to code for what we consider to be the main function of each token.

Focus
This term is not particularly specific, but is found throughout the literature.

Miller and Weinert describe *like* as a non-contrastive focusing device (1995: 365). Underhill (1988) suggests it focuses attention on the following information. Dailey-O'Cain labels focuser *like* as a discourse or pragmatic marker, saying that it can be used 'to initiate, sustain, or repair discourse, to mark a boundary or sequential dependence between discourse segments, or to denote either new or old information in informal speech' (2000: 61). It can also be used for the 'elucidation of previous comments' (Miller and Weinert 1995: 366).

(1) *get to a-maybe a college like after I do the Highers* (EE013, 00.08.42.63)[3]

(2) *as in like 'I know'* (EE003, 00.01.19.31)

(3) *we speak like all the time with them* (EP003, 00.05.03.42)

Word-finding difficulties
This is when the speaker is struggling to find the right word and uses *like* to fill the thinking pause. This includes cases such as (4)–(5) where there are clear prosodic hesitations (shown with full stop punctuation, see also (8)) or cases such as (6)–(8) where the context clearly shows the speaker is searching for the most appropriate word (and where *like* may serve a similar function as a filled pause such as *um*) or where the speaker seems to be 'backstepping' (Craig and Sanusi 2000: 433), that is revising something that has just been proffered (9).

(4) *but she's like. skinny and chubby if you know what I mean* (EE008, 00.22.44.94)

(5) *that's like your. your eh, no* (EE011, 00.15.39.44)

(6) *the other things are from. like, I don't know, ASDA is our nearest* (EP010, 00.31.07.55)

(7) *I want to like. I don't know, maybe go-go to college* (EP015, 00.07.50.38)

(8) *and on like. you know MSN* (EP009, 00.02.17.97)

(9) *I think they're more likely to pronoun-like say 'water'* (EE003, 00.07.54.08)

Since the these functions can be particularly hard to tease apart, they have been grouped together and coded accordingly. Miller says that *like* 'does not occur at pauses or where the speaker has problems planning the syntax' (2003: 108), but as can be seen, we did find some instances where this does appear to be the case in our data.

Quotative

This appears in the form BE + *like* and acts as a 'verbal element, inflectable for both tense and agreement' (D'Arcy 2005: 3). In this case, *like* often serves a mimetic function – highlighting the performative dimensions of the reported discourse (Buchstaller 2003, Levey 2003).

(10) *cos you're like 'where am I fae', instead of 'from'* (EE006, 00.04.32.37)

Approximation

Like may also be used as a means of expressing approximation to a measure or quantity, e.g. 'I was out of school for *like* seven months'. Following D'Arcy (2005), who argues that when *like* signals approximation it is an adverbial and not a DM, we have excluded tokens with approximation functions from the results reported below. We note, however, that Buchstaller (2014) makes a convincing theoretical case for a chain from the use of *like* as an approximation to *(be) like* as a quotative, and backs this up with some telling early examples in her corpora (2014: 154). It may be that notwithstanding D'Arcy's arguments, there is a still a case for including approximation functions of *like* in a broader survey of its distribution. In the present study, it makes little difference: with or without the approximation tokens, the same independent factors are selected as significant for the Edinburgh- and Polish-born sub-corpora.[4]

Clause-final modifier

Miller and Weinert state that clause-final *like* is 'used to (anticipate and) counter (possible) objections and assumptions' (1995: 23) (though whether there is a categorical or probabilistic association between position and function is an empirical question). An example observed in the data is 'or something like that *like'*. Since this is a common usage in Scottish (and other Northern) English (Nestor 2013), it is important to retain this context.

    (11) *depending on who I'm mucking about with that day like* (EE009, 00.13.02.48)

    (12) *oh day ticket, like* (EP003, 00.18.16.77)

Corrective

A rephrasing or reformulation function to *like* is also mentioned in the literature: 'In certain cases, *like* appears to frame a restart where a speaker starts off on a particular track and then feels the need to rephrase' (Levey 2003: 27).

    (13) *so they mo- like she moved down here, so did my auntie and stuff* (EE008, 00.18.16.77)

    (14) *but I still like- I feel embarrassed or something* (EP010, 00.05.57.29)

Unclear/other

This code was used where it was unclear as to which function was being used, or where a different sort of function than the ones listed above appeared to be being used. There are only twenty-three tokens with *Unclear/Other* functions produced by the Edinburgh- and Polish-born teenagers in total.

3.3 Independent variables

Grammatical context

For each token, the preceding and following grammatical context was determined. This included: other discourse markers (e.g. '<u>Just</u> *like*'), noun phrases (e.g. '*like* <u>Scottish people</u>'), verb phrases (e.g. '*like* <u>it all just works</u>'), prepositional phrases (e.g. '*like* <u>in October</u>'), complementisers (e.g. 'maybe *like* <u>that she wants to</u>'), adjectives (e.g. 'when I was *like* <u>one</u>'), pauses (e.g. 'nothing really <u>…</u> *like*'), conjunctions (e.g. '<u>but</u> *like* yeah'), pronominals (e.g. '<u>he</u> *like* mumbles a lot'), negative markers (e.g. '<u>not</u> *like* close friends'), auxiliaries (e.g. 'he<u>'s</u> *like*', 'but that<u>'s</u> *like* if it's a nice weather'), verbal pauses (instead of *um* and *ah*), adverbials ('play tennis, *like* <u>randomly</u>'), within relative clauses ('there was one group <u>who was</u> *like*'), and instances where it was unclear as to which grammatical category the preceding or following context belonged, as in (15).

    (15) *I'm not joking but they like [sound effects]* (EE010, 00.03.16.21)

A note on JUST

Just ordinarily acts as an adverb, but it is coded as 'other discourse function' since, in this data it appears to be acting with *like* as a focuser or hedge. That is, the two frequently appear together and seem to act as a pragmatic expression of either 'the speaker's uncertainty concerning the choice of the following word or phrase' (Holmes 1986: 10), as in 'theirs is just *like*… different' or to enhance the focusing nature of *like*, as in 'oh they're just *like* together'.

Non-linguistic (social) factors

It has been argued that gender is a socially salient and highly 'available' category for variation to be mapped onto (Meyerhoff & Schleef 2011) even for non-native speakers. Coding for gender allows us to test this claim.

We also coded the Polish speakers for three non-linguistic measures that we thought might influence their chances of acquiring local-sounding norms for the use of *like*. We differentiated between migrant teenagers whose friendship networks seemed (based on what they said in the interviews) to be mainly Polish, and contrasted this with Polish kids whose friends were mainly Scottish or who seemed to have a very mixed friendship network.

We coded the Polish teenagers also for the amount of time they reported that they had been formally learning English distinguishing those who had learnt it for more than five years and those who had learnt it less.

Finally, since most of the functions of *like* are not taught in English classes, we coded the Polish kids for how long they had been in Scotland, differentiating those who had been there for more than five years from the rest. In the course of our analysis, we found that formal learning and informal exposure seemed to be interacting with each other. As a consequence, we made a composite measure of 'Exposure to English' which captures the Polish teenagers' formal and informal exposure to the language.

Friendship network was not entirely independent of both the learning factors. Those with mainly Scottish networks almost always were the ones who had been (formally) learning English longer and who had been in Scotland longer. However, friendship network appears to be operating independently of the other measures of English – perhaps unsurprisingly since it may encapsulate attitudinal factors that are opaque with the other temporal measures.

4. Findings

4.1 Frequency of like *variants*

The Edinburgh- and Poland-born teenagers used *like* with very different frequencies overall, but proportionally the distribution of *like* with the six main discourse functions we coded for was very similar across the two groups.

Table 1: Distribution of tokens of *like* across the different functions (as percentage of group total)

	FOCUS	WORD FINDING	CLAUSE-FINAL	CORRECTIVE	QUOTATIVE	UNCLEAR/ OTHER	TOTAL
Edinburgh-born	401 (68%)	86 (15)	24 (4)	20 (3)	48 (8)	10 (2)	589
Poland-born	225 (69)	55 (17)	5 (2)	10 (3)	15 (5)	13 (4)	323
Total	626	141	29	30	63	23	912

The multivariate analysis did not find the distinction between native and non-native speaker to be a significant factor group in predicting the use of focus *like* versus *like* with any other function. This suggests that despite their difficulties with English overall (some of the speakers in this sample were still very hesitant), the Polish teenagers are starting to acquire the use of *like* as a DM.

However, if we use the distinction between native and non-native speaker as the dependent variable, the function of *like* does prove to be a significant predictor. This suggests that there are some important differences in how the Polish teenagers are using the less frequent functions of *like.* For example, they are less likely to use *like* as a quotative and in clause-final position than the locally-born teenagers are. The Polish teens are slightly more likely to use it in contexts when there is some kind of word finding problem than the native speakers are, but this is likely to be related to the simple fact that they are non-native speakers and because they are all at rather early stages in their learning of English are more likely to have problems with lexical retrieval in general and may well appreciate the usefulness of this function of *like*.

As Table 1 shows, the predominant function for *like* for both groups of speakers is as a focus marker (68% and 69% of all total utterances of *like*).

We have, therefore, taken this as the application value for all subsequent analyses. We are essentially asking whether any of the other linguistic and non-linguistic factor groups significantly constrain the use of *like* in its main function (as a focuser) as opposed to its use in all other discourse functions.

There are of course considerable interspeaker differences in the frequency with which *like* is used. Rbrul allows us to treat individual speaker as a random effect, which means the model is less likely to overestimate the effect of any of the other linguistic and non-linguistic factors included.

4.2 Constraints on the functions of *like*

In the following sections, we present the results for the factors that were found to be significant constraints on the use of *like* in the corpus. Where the factor group was significant for both the Edinburgh- and Poland-born teenagers, we present the ranking of constraints in a way that enables some comparison across the two groups.

The following constituent proved to be the strongest fixed effect for the function that both groups of teenagers use *like* to express. Tables 2 and 3 show the frequency and probability of *like* occurring as a focus marker versus all other functions in the two sub-corpora (the NS and NNS groups were analysed separately and we have separated out the results for each of the statistically significant effects in the tables that follow simply in the interests of clarity of presentation, i.e. Tables 2 and 4 come from the same model).

For the Edinburgh-born teenagers, the only other significant constraint on the use of *like* as a focus marker was the nature of the preceding constituent. The results are shown in Table 4.

For the Polish teenagers, the preceding constituent was not a significant constraint on *like* as a focuser. One non-linguistic factor was a significant predictor: speaker sex/gender (Table 5).

From Table 5 we can see that *like* as a focuser is used by the Polish girls more whereas the Polish boys are more likely to use it more in its other functions. This suggests that the teenage girls are approaching the norms of their Edinburgh-born peers, perhaps ahead of the boys. Recall that gender is not a significant factor for the Edinburgh-born teenagers. Although the Edinburgh-born girls use *like* much more often than the boys do overall (75% to 25% total tokens of *like*)[5], both the locally-born girls and boys use *like* with the same functions. That is, among the girls, *like* serves as a focuser 64% of the time they use it as a DM; among the boys it serves as a focuser 62% of the time they use it.

Table 2: Significant constraints on *like* for native speakers - following context

NATIVE SPEAKERS				
FOLLOWING CONTEXT	LOGODDS	TOKENS	% FOCUS MARKER	FACTOR WEIGHT
adjective	1.792	33	0.909	0.857
pronoun	1.751	152	0.875	0.852
prepositional phrase	1.726	20	0.900	0.849
verb	0.621	238	0.748	0.650
relative clause	-0.097	2	0.500	0.476
adverbial	-1.132	14	0.357	0.244
other discourse marker	-1.327	5	0.400	0.210
pause	-1.666	120	0.267	0.159
conjunction	-1.667	3	0.333	0.159
Total		587		

Table 3: Significant constraints on *like* for non-native speakers – following context

NON-NATIVE SPEAKERS				
FOLLOWING CONTEXT	LOGODDS	TOKENS	% FOCUS MARKER	FACTOR WEIGHT
adjective	1.620	28	0.893	0.835
prepositional phrase	0.935	15	0.867	0.718
pronoun	0.860	94	0.830	0.703
verb	0.166	118	0.737	0.541
adverbial	-0.708	5	0.600	0.330
other discourse marker	-1.029	4	0.500	0.263
pause	-1.844	58	0.293	0.137
Total		322		

Table 4: Significant constraints on *like* for native speakers – preceding context

NATIVE SPEAKERS				
FOLLOWING CONTEXT	LOGODDS	TOKENS	% FOCUS MARKER	FACTOR WEIGHT
adjective	1.432	10	0.800	0.807
conjunction	0.679	35	0.743	0.664
pause	0.449	101	0.832	0.610
negative marker	0.327	8	0.875	0.581
adverbial	0.271	25	0.680	0.567
other discourse marker	0.176	45	0.711	0.544
prepositional phrase	0.157	34	0.824	0.539
auxiliary	-0.205	79	0.709	0.449
nominal	-0.272	27	0.593	0.432
pronominal	-0.866	18	0.611	0.296
verbal pause	-1.063	5	0.400	0.257
verb	-1.085	200	0.565	0.253
Total				

Table 5: Female/male differences for non-native speakers

NON-NATIVE SPEAKERS				
FOLLOWING CONTEXT	LOGODDS	TOKENS	% FOCUS MARKER	FACTOR WEIGHT
Female	0.364	252	0.714	0.59
Male	-0.364	70	0.643	0.41

This finding seems to be consistent with a proposal put forward in Meyerhoff & Schleef (2011). That paper discusses the emergence of a significant gender effect among the Polish teenagers in their use of the variable (ing), where no such effect is found in the speech of the locally-born peers. Meyerhoff & Schleef (2011) argue that gender is a socially salient category (meaning it is a category that is readily available for comment and is already associated with probabilistic differences in speech patterns), and therefore when the Polish learners of English may be working hard to systematise and

replicate the variation heard around them, gender provides an accessible and retrievable scaffold on which to hang a new variable.[6]

4.4 Time spent in an English-speaking environment

The results showed that time spent in the UK did not have a significant effect, but time spent learning English did. Interestingly, males who had been learning English for a shorter period of time (two – five years as opposed to over five years) did not use quotative *like* at all. However, the females used *like* across all functions, no matter how long they had been learning English.

5. Conclusion

At the beginning of this study we asked three questions:

1. Do the Polish-born teenagers in the corpus use the DM *like*?

2. If they do,
    a. do they use it with the same frequency as their locally-born peers?
    b. do they use it for the same discourse functions?
    c. is the variation in their use of *like* subject to the same social and linguistic constraints as the variation in locally-born speakers' use of *like*?

It is clear that the answer to the first question is yes. However, the frequency of use of *like* by the non-native speakers is not as high as the native speakers: there are 323 total utterances of *like* by the Polish kids, compared to 589 by their Scottish peers. Miller and Weinert note that *like* has 'a role which is learned relatively late, apparently after age ten, by native-speaker learners' (1995: 366). This suggests that the Polish speakers may well expand their functional range of *like* quickly, since the native speakers themselves have only been using *like* for a short period of time.

In terms of the discourse function of utterances of *like* from our corpus, the following grammatical constituent proved to be the strongest constraint for both groups of speakers. For the native speakers, the only other significant constraint on the use of *like* as a focus marker was the nature of the preceding constituent.

For the Polish teenagers, the preceding constituent was not a significant

constraint on *like* as a focuser. However, one non-linguistic factor was a significant predictor: speaker gender. We found that *like* as a focuser is used by the Polish girls more whereas the Polish boys are more likely to use it more in its other functions suggesting that the teenage girls are moving towards the norms of their Edinburgh-born peers ahead of the boys. This was not a significant factor for the native speakers, with the girls using *like* as a focuser 64% of the time they use it as a DM; while the boys use it as a focuser 62% of the time.

We also considered which of the following possible outcomes for the transfer of variation in the use of DMs in the speech of one group of speakers to the speech of another group of speakers might occur:

1. No transfer (failure to acquire anything like the L1 patterns of variation);

2. Transfer of the L1 form but failure to acquire the L1 constraints on the variation;

3. Transfer of both the L1 form and the L1 constraints.

There has been transfer of both form and function, though the latter has not been transferred to the same extent as the former. As discussed above, the Polish-born teens are catching up to their Edinburgh-born peers in terms of the spread of their use of *like* across the different discourse functions.

Notes

1 We thank the UK Economic and Social Research Council for support on the project 'Sociolinguistics of Immigration', (PI: M. Meyerhoff; Co-I: E. Schleef). We also thank two anonymous *Te Reo* reviewers for their constructive feedback that has helped improve the paper. Infelicities in the text are entirely our responsibility.

2 Given the total number of tokens reported for other XPs, it seems that this sample is also the basis for the analysis of *like* in DPs/NPs, while for VPs it seems only the declarative clauses in this sub-sample were analysed.

3 EE = Edinburgh-born Edinburgh teenage; EP= Polish-born Edinburgh teenager. The numbers refer to speaker number in our corpus followed by time stamp.

4 There are some minor changes to the ranking of factors within factor groups depending on whether approximation tokens (seventy-four tokens total across the two groups) are included or not.

5 A figure that needs to be treated with caution since we have not quantified over total number of words.

6 It is worth noting that gender appears to interact with the exposure to English and friendship network factors as well. A first pass over the data for the Polish teenagers indicated that exposure to English and friendship network were significant factors (and not gender). However, this effect disappeared after we recoded some of the verbal and nominal factors in the following constituent group together to eliminate categorical tokens.

References

Bazzanella, Carla, Cristina Bosco, Alesandro Garcea and B. Gili Fivela 2007. 'Italian *allora*, French *alors*: Functions, convergences and divergences'. *Catalan Journal of Linguistics 6*: 9–30.

Blakemore, Diane 2002. *Relevance and Linguistic Meaning: The semantics and pragmatics of discourse markers*. Cambridge: Cambridge University Press.

Brinton, Laurel J. 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin and New York: Mouton de Gruyter.

Buchstaller, Isabelle 2003. 'The co-occurrence of quotatives with mimetic performances'. Edinburgh Working Papers in Applied Linguistics 12: 1–9.

Buchstaller, Isabelle 2014. *Quotatives: New Trends and Sociolinguistic Implications.* Oxford: Wiley-Blackwell.

Cedergren, Henrietta and David Sankoff 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50: 333–355.

Cheshire, Jenny 2005. 'Syntactic variation and beyond: gender and social class variation in the ide of discourse-new markers'. *Journal of Sociolinguistics* 9: 479–507.

Cheshire, Jenny 2007. 'Discourse variation, grammaticalisation and stuff like that'. *Journal of Sociolinguistics* 11: 155–193.

Clift, Rebecca 2001. 'Meaning in interaction: The case of *actually*'. *Language* 77: 245–291.

Craig, Robert T. and Alena L. Sanusi 2000. ' 'I'm Just Saying . . .': Discourse Markers of Standpoint Continuity'. *Argumentation* 14: 425–445.

Cresswell, M. J. 1990. *Entities and Indices*. Dordrecht, Boston & London: Kluwer Academic Publishers.

D'Arcy, Alexandra F. 2005. *Like: Syntax and Development.* Unpublished PhD, University of Toronto.

Dailey-O'Cain, Jennifer 2000. 'The sociolinguistic distribution of and attitudes towards focuser *like* and quotative *like*'. *Journal of Sociolinguistics* 4(1): 60–80.

Dines, Elizabeth R. 1980. 'Variation in discourse – "and stuff like that"'. *Language in Society* 9: 13–33.

Fraser, Bruce 1990. 'An approach to discourse markers'. *Journal of Pragmatics* 14: 383–395.

Fraser, Bruce 2009. 'An account of discourse markers'. *International Review of Pragmatics* 1: 293–320.

Hansen, Maj-Britt M. 1995. *'Puis* in Spoken French: from time-adjunct to additive conjunct?' *French Language Studies* 5: 31–56.

Hansen, Maj-Britt M. 1998. 'The semantic status of discourse markers'. *Lingua* 104: 235–260.

Hansen, Maj-Britt M. 2005. 'A comparative study of the semantics and pragmatics of enfin and finalement, in synchrony and diachrony'. *French Language Studies* 15: 153–171.

Heritage, John 1998. '*Oh*-prefaced responses to inquiry'. *Language in Society* 27: 291–334.

Heritage, John 2013. 'Turn-initial position and some of its occupants'. *Journal of Pragmatics* 57: 331–337.

Holmes, Janet 1986. 'Functions of *you know* in women's and men's speech'. *Language in Society* 15: 1–22.

Holmes, Janet 1988. '*Sort of* in New Zealand women's and men's speech'. *Studia Linguistica* 42: 85–121.

Holmes, Janet 1989. 'Review of Bent Preisler, Linguistic sex roles in conversation'. *Language in Society* 18: 293–299.

Hopper, Paul J. and Elizabeth C. Traugott 2009. *Grammaticalization* (2nd ed.). Cambridge: Cambridge University Press.

Johnson, Daniel E. 2009. 'Getting off the Goldvarb Standard: Introducing Rbrul for mixed-effects variable rule analysis'. *Language and Linguistics Compass* 3: 359–383.

Kearns, Kate 2000. *Semantics*. London: Longman.

Labov, William 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Lakoff, George 1973. 'Hedges: A study in meaning criteria and the logic of fuzzy concepts'. *Journal of Philosophical Logic* 2: 458–508.

Levey, Stephen 2003. 'He's like 'Do it now!' and I'm like 'No!' '. *English Today, 19*(1): 24–32.

Matras, Yaron 2012. 'An activity-oriented approach to contact-induced language change'. In C. Chamoreau and I. Léglise (eds.) *Dynamics of Contact-Induced Change*, Berlin: Mouton de Gruyter, 17–52.

Meyerhoff, Miriam 1994. ' "Sounds pretty ethnic, eh?" – a pragmatic particle in New Zealand English'. *Language in Society*. 23, 3. 367-388.

Meyerhoff, Miriam 2009. 'Replication, transfer and calquing: Using variation as a tool in the study of language contact'. *Language Variation and Change*. 21: 1–21.

Meyerhoff, Miriam and Erik Schleef 2012. 'Variation, contact and social indexicality in the acquisition of (ing) by teenage migrants'. *Journal of Sociolinguistics*, 16: 398–416.

Miller, Jim 2003. 'Syntax and Discourse in Modern Scots'. In J. Corbett, J. Derrick McClure and J. Stuart-Smith (Eds.), *The Edinburgh Companion to Scots*. Edinburgh: Edinburgh University Press. 72–109.

Miller, Jim and Regina Weinert 1995. 'The function of LIKE in dialogue'. *Journal of Pragmatics 23*: 365–393.

Nestor, Niamh 2013. 'The positional distribution of discourse *like*: a case study of young Poles in Ireland'. In David Singelton, Vera Regan and Ewelina Debaene (eds.) *Linguistic and Cultural Acquisition in a Migrant Community*. Bristol: Multilingual Matters. 49–72.

Paolillo, John 2013. 'Individual effects in variation analysis: Model, software, and research design'. *Language Variation and Change* 25: 89–119.

Pichler, Heike 2010. 'Methods in discourse variation analysis: Reflections on the way forward'. *Journal of Sociolinguistics* 14: 581–608.

Preisler, Bent 1986. *Linguistic Sex Roles in Conversation*. Berlin: Mouton de Gruyter.

Rini, A. A. and M. J. Cresswell 2012.*The World-Time Parallel*. Cambridge: Cambridge University Press.

Rosch, Eleanor H. 1973. 'On the internal structure of perceptual and semantic categories'. In T. E. Moore (Ed.). Cognitive development and the acquisition of language. New York: Academic Press. 111–144.

Rosch, Eleanor H. and C.B. Mervis 1975. 'Family resemblances: Studies in the internal structure of categories'. *Cognitive Psychology* 7: 573–605.

Sankoff, Gillian and Pierette Thibault 1977. 'L'alternance entre les auxiliaires *avoir* et être en français parlé à Montréal›. *Langue française* 34: 84–108. Also published in English in Sankoff, G. 1980. *The Social Life of Language.* Philadelphia: University of Pennsylvania Press.

Sankoff, Gillian, Pierette Thibault, Naomi Nagy, Hélène Blondeau, Marie-Odile Fonollosa and Lucie Gagnon 1997. 'Variation in the use of discourse markers in a language contact situation'. *Language Variation and Change* 9: 191–217.

Schegloff, Emmanuel A. 1988. 'Discourse as an interactional achievement, II: An exercise in conversation analysis'. In D. Tannen (ed.) *Linguistics in context: Connecting observation and understanding: Lectures from the 1985 LSA/TESOL and NEH Institutes.* Norwood, NJ: Ablex. 135–58.

Schegloff, Emmanuel A. 1993. 'Reflections on Quantification in the Study of Conversation'. *Research on Language and Social Interaction* 26: 99–128.

Schiffrin, Deborah 1987. *Discourse markers.* Cambridge & New York: Cambridge University Press.

Schleef, Erik, Meyerhoff, Miriam and Clark, Lynn 2011. 'Teenagers' acquisition of variation: A comparison of locally-born and migrant teens' realisation of English (ing) in Edinburgh and London'. *English World-wide 32*(2): 206–236.

Schourup, Lawrence C. 1985. *Common Discourse Particles in English Conversation*. New York: Garland.

Schourup, Lawrence C. 1999. 'Discourse Markers'. *Lingua* 107: 227–265.

Tannen, Deborah (ed.) 1988. *Linguistics in context: Connecting observation and understanding: Lectures from the 1985 LSA/TESOL and NEH Institutes.* Norwood, NJ: Ablex.

Underhill, Robert 1988. 'Like is like, focus'. *American Speech* 63: 234–246.

Walker, James A. 2010. *Variation in Linguistic Systems*. London: Routledge.

Wohlgemuth, Tracy L. 2003. *It's like discourse 'like': The story of 'like' in Canadian English in 1995*. Unpublished MA paper, University of Toronto.

Wouk, Fay 1999. 'Gender and the use of rpagmatic particles in Indonesian'. *Journal of Sociolinguistics* 3: 194–219.

Zadeh, Lofti A. 1965. 'Fuzzy sets'. *Information and Control* 8: 338–353.

CATEGORY AMBIGUOUS WORDS COMPARED ACROSS THREE TYPES OF CORPORA

Elaine W. Vine: *School of Linguistics and Applied Language Studies, Victoria University of Wellington <elaine.vine@vuw.ac.nz>*

Abstract

This study draws on three types of corpora: general English corpora, learner English corpora and a corpus of English language teaching coursebooks, in an investigation of high frequency words which are category ambiguous, that is, one word-form has different grammatical uses. Four words are analysed and discussed: *about*, *over*, *end* and *place*. Variation in patterns of use of these words is shown within and across the corpora, which gives rise to discussion of pedagogical implications. Patterns of use in coursebooks often differ from patterns of use in general English, and it is argued that while corpus frequency data should not necessarily be pedagogically prescriptive, they should inform pedagogy except where there is good reason for them not to do so.

1. Introduction

Over the past decade, there has been significant progress in the development of 1) general English corpora (e.g. O'Keeffe et al. 2007) and 2) learner corpora (e.g. Granger 2009) and their application to English language teaching. There has also been study of English language use in 3) coursebook corpora (e.g. Meunier & Gouverneur 2009). This work has built on the insights of, for

example, Kennedy (1987) who argued in his study of how quantification is expressed in written English that corpus-based information about patterns of use 'surely is an improvement on using impressionistic means alone as a basis for English language teaching' (p. 283).

Many words in English, particularly words that are used frequently, are multifunctional. As Leung (1991) noted in his study of the meanings of *over*, this poses challenges to learners of English, and to their teachers, who have to decide when and how to teach the words and their various uses. However, Leung's study was not corpus-based, so he could not comment on how frequently each of the meanings of *over* that he identified is used in English.

If, in English language teaching contexts, we are to apply Kennedy's suggestion about corpus-based information and Leung's insight about multifunctional words, it is relevant to first explore patterns of use across different types of corpora. In this paper, I draw on the three types of corpora mentioned above to focus on category ambiguity in high frequency words such as *over*, that is, where one word-form has different grammatical uses, for example, *over* as adverb (i) and *over* as preposition (ii):

 i. i'll come **over** and stay a few days (dpc192)

 ii. there is a row in the government **over** the report (msn062)
      (examples from *Wellington Corpus of Spoken New Zealand English*).

Other main types of lexical ambiguity do not involve grammatical differences: for so-called 'within-category homonymy', unrelated words share the same word-form, for example, *bat* as a flying mammal and as a piece of sports equipment, and for polysemy a single word has related senses, such as *point* as the tip of a blade or as a promontory.

My research questions are:

1. What are the patterns and frequencies of use of high frequency category-ambiguous words in general English corpora, learner corpora and English language coursebooks?

2. To what extent are the patterns and frequencies of use of these words in general and learner English corpora reflected in pedagogical applications in coursebooks?

In this paper, I have chosen to focus on *over* first, to provide a corpus-based perspective on that word which was missing from Leung's (1991) study. I have

also chosen three other words to compare with *over*: *about*, because, like *over*, it has adverb (iii) and preposition (iv) uses:

   iii. i stayed there for **about** an hour afterwards (dpc036)

   iv. quentin was thinking **about** becoming a d j (dpf024)

The other two words are *end* and *place*. These were chosen to contrast with *over* and *about* in that, while teachers regard *over* and *about* as 'function words', they regard *end* and *place* as 'content words'. *End* and *place* have both noun (v and vii) and lexical verb (vi and viii) uses:

   v. you did all walk out at the **end** just as I was trying to get one question resolved (dpc308)

   vi. i suppose it's quite hard to **end** a movie like that though (dpc273)

   vii. it's a nice **place** for a couple (dpc008)

  viii. that's the only er view i could **place** on it (dgi148)

All four words are high frequency words in English, as indicated by their frequency rankings in an unlemmatised list of word-forms prepared from the BNC for use by English learners (Audience Dialogue 2006): *about* (55th), *over* (74th), *end* (169th) and *place* (173rd).

I report here on analyses and comparisons of patterns of use of these four category ambiguous words in general English corpora, learner corpora, and a coursebook corpus.

2. Comparison corpora

The comparison corpora for this study are:

   General English
      British English (BE)
         Spoken BE – spoken section of the British National Corpus
         Written BE – written section of the British National Corpus

The British National Corpus (BNC) was completed in 1994. The spoken section contains *c.* 10 million words and the written section *c.* 90 million words. For this study the BNC was searched using the on-line interface *BYU-*

BNC: The British National Corpus (Davies 2004). All other corpora were searched using *WordSmith Tools* (Scott 2004).

> New Zealand English (NZE)
>> Spoken NZE – Wellington Corpus of Spoken New Zealand English
>> Written NZE – Wellington Corpus of Written New Zealand English

The two Wellington corpora contain 1 million words each. The spoken corpus is made up of texts collected between 1988 and 1994 (Holmes *et al.* 1998) and the written corpus is made up of writings published between 1986 and 1990 (Bauer 1993).

> *L1 Learner English (L1 LE)*
>> Spoken L1 LE – Louvain Corpus of Native English Conversation (LOCNEC) – learner turns only

The learner turns section of LOCNEC contains 124,352 words. It is composed of interviews with British university undergraduate students. It is a comparison corpus for LINDSEI (see below) and has the same format and tasks.

> Written L1 LE – Louvain Corpus of Native English Essays (LOCNESS)

The LOCNESS corpus contains 288,177 words: 149,574 from argumentative essays by American university students, 18,826 from literary-mixed essays by American university students, 59,568 from argumentative and literary essays by British university students, and 60,209 from British A-level argumentative essays. It is a comparison corpus for ICLEv2 (see below).

> *L2 Learner English (L2 LE)*
>> Spoken L2 LE – Louvain International Database of Spoken English Interlanguage – learner turns only (LINDSEI: Gilquin *et al.* 2010)

LINDSEI is a 1 million word corpus of 554 interviews with undergraduate university students who are higher intermediate to advanced EFL learners, most from countries in Europe, with some from China and Japan. The learner turns within the interviews total just under 800,000 words.

> Written L2 LE – International Corpus of Learner English (ICLEv2: Granger *et al.* 2009)

ICLEv2 is a 3.75 million word corpus of undergraduate university writing (mostly argumentative essays) by higher intermediate to advanced EFL

learners, most from countries in Europe, with some from China, Japan and South Africa.

English language teaching coursebook series
Cutting Edge (CE)
Starter (1) (Cunningham, Redston & Moor 2007), Elementary (2) (Cunningham, Moor & Eales 2005), Pre-Intermediate (3) (Cunningham, Moor & Carr 2005), Intermediate (4) (Cunningham & Moor 2005), Upper Intermediate (5) (Cunningham & Moor 2007), Advanced (6) (Cunningham, Moor & Carr 2007)

New Headway (NH)
Beginner (1) (Soars & Soars 2002), Elementary (2) (Soars *et al.* 2000), Pre-Intermediate (3) (Soars *et al.* 2007), Intermediate (4) (Soars & Soars 2003b), Upper-Intermediate (5) (Soars & Soars 2005), Advanced (6) (Soars & Soars 2003a)

The coursebook corpus comprises the full texts of all student books in the two series of coursebooks, including contents pages; activities and exercises; written texts; transcripts of spoken texts, words on pictures, diagrams, charts; answer keys. The two series are widely used with adult learners of English, ranging from beginners through to advanced learners.

Table 1 shows the number of words in each book of each of the two coursebook series. Some differences can be seen between books at each level in the two series, but the total number of words for each series is similar.

Table 1: Coursebook corpus: CE and NH series

CUTTING EDGE (CE)	NO. OF WORDS	NEW HEADWAY (NH)	NO. OF WORDS
CE 1	35,209	NH 1	33,773
CE 2	70,102	NH 2	57,443
CE 3	81,231	NH 3	79,318
CE 4	90,306	NH 4	78,537
CE 5	96,435	NH 5	99,458
CE 6	88,097	NH 6	96,510
Total	461,380		445,039

While the comparison corpora vary considerably in size, even the smaller ones are adequate for the purposes of this study given that the four target words are all high frequency words.

3. Comparison findings

All of the analyses presented in this paper relate to exact forms of the relevant words (e.g. *end*), not inflected forms (e.g. *ends*, *ended*, *ending*) because some of the inflected forms (e.g. *ended*) are not category ambiguous. In the following tables and figures, frequency data are presented in standardised form per 10,000 words. However, the log-likelihood (LL) statistics were calculated from the raw data, using the online log-likelihood calculator at http://ucrel.lancs.ac.uk/llwizard.html which was developed by the University Centre for Computer Corpus Research on Language at Lancaster University.

3.1 General English corpora

As shown in Table 2, there are statistically significant differences between the two spoken corpora for all four words, but *about*, *over* and *place* occur more frequently in NZE than in BE, while *end* occurs more frequently in BE than NZE.

Table 2: Spoken BE and NZE: occurrences per 10k words

	about	*over*	*end*	*place*
BE – Spoken	36	10	5.4	3.9
NZE – Spoken	40	13	4.6	5.6
	LL 46.96 $p<0.0001$	LL 64.86 $p<0.0001$	LL 10.14 $p<0.01$	LL 61.90 $p<0.0001$

As shown in Table 3, in the written corpora, three words, *about*, *over* and *place*, occur more frequently in NZE than in BE, and one word, *end*, occurs more frequently in BE than in NZE. but the differences are only statistically significant for *about*, *over* and *end*.

Table 3: Written BE and NZE: occurrences per 10k words

	about	*over*	*end*	*place*
BE – Written	18	14	4.8	4.9
NZE – Written	20	15	4.3	5.2
	LL 38.17 p<0.0001	LL 5.20 p<0.05	LL 8.10 p<0.01	LL 0.54 ns

3.2 Learner English corpora

3.2.1 L1 learner English corpora

The data for L1 learner English (L1 LE) (see Table 4) show that *about*, *end* and *place* occur more frequently in spoken L1 LE than in written L1 LE. By contrast, *over* occurs at the same frequency in both corpora.

Table 4: Spoken and written L1 LE: occurrences per 10k words

	about	*over*	*end*	*place*
L1 LE – Spoken	40	10	9.4	6.8
L1 LE – Written	18	10	6.3	5.0
	LL 153.96 p<0.0001	LL 0.00 ns	LL 10.55 p<0.01	LL 4.66 p<0.05

3.2.2 L2 learner English corpora

The data for L2 learner English (L2 LE) (see Table 5) show that *about* occurs more frequently in spoken L2 LE than in written L2 LE. By contrast, the other three words occur more frequently in written L2 LE than in spoken L2 LE, though the difference for *end* is not significant.

Table 5: Spoken and written L2 LE: occurrences per 10k words

	about	*over*	*end*	*place*
L2 LE – Spoken	37	3.6	5.1	6.4
L2 LE – Written	27	6.3	5.2	8.0
	LL 210.91 p<0.0001	LL 89.70 p<0.0001	LL 0.01 ns	LL 23.92 p<0.0001

3.3 L1 and L2 learner English corpora compared

3.3.1 Spoken L1 and L2 learner English corpora

In the parallel learner spoken corpora, L1 learners use all four words more frequently than L2 learners, but the differences are statistically significant only for *over* and *end* (see Table 6).

Table 6: Spoken L1 LE and L2 LE: occurrences per 10k words

	about	*over*	*end*	*place*
Spoken – L1 LE	40	10	9.4	6.8
Spoken – L2 LE	37	3.6	5.1	6.4
	LL 1.97 ns	LL 78.76 p<0.0001	LL 28.16 p<0.0001	LL 0.20 ns

3.3.2 Written L1 and L2 learner English corpora

The pattern is different for the parallel learner written corpora. Here, L1 learners use *over* and *end* more frequently than L2 learners, but the reverse is the case for *about* and *place*. In all four cases the differences are statistically significant (see Table 7).

Table 7: Written L1 LE and L2 LE: occurrences per 10k words

	about	*over*	*end*	*place*
Written – L1 LE	18	10	6.3	5.0
Written – L2 LE	27	6.3	5.2	8.0
	LL 102.30 p<0.0001	LL 63.21 p<0.0001	LL 7.39 p<0.01	LL 39.94 p<0.0001

3.4 Coursebook corpora

One of the target words, *about*, occurs significantly more frequently in the CE series of coursebooks than in the NH series (see Table 8). The differences for the other three words are not statistically significant.

There are also differences between the two series on a book-by-book basis (see below).

Table 8: Target words in CE and NH series per 10k words

	about	*over*	*end*	*place*
CE	52	8.2	4.4	6.3
NH	44	8.5	5.0	5.8
	LL 32.30 p<0.0001	LL 0.21 ns	LL 1.59 ns	LL 1.05 ns

3.5 General English, L2 and L1 learner English and coursebook corpora compared

This section includes calculations of frequencies of use of the target words according to the grammatical category of use of each occurrence of the word. For the BNC, the standardised data were calculated by BYU-BNC (Davies 2004) on the basis of the automatic part-of-speech tagging in the corpus. For all other corpora, the standardised data for grammatical categories were calculated on the basis of frequencies extrapolated from a manual analysis of a random sample, generated by *WordSmith Tools*, of 100 occurrences of the target word from throughout each corpus. A sample size of 100 occurrences from each corpus was chosen to be large enough to provide sufficient evidence of different grammatical uses of the target word, while also being manageable for manual analysis. Significance of findings has been calculated in SPSS using Chi square for comparing categorical data, assuming expected equal frequencies.

3.5.1 about

In the case of *about* (see Figure 1), the overall pattern noted earlier that *about* occurs more frequently in CE than in NH (see Table 8) is replicated in each of the six books in the series. There is also a trend in both series for the frequency of *about* to decrease across the series, though that trend is not evenly spread across the series.

However, Figure 1 also shows that *about* is used more frequently in coursebooks than in spoken English from all corpora, and much more frequently than in written English from all corpora. Also, written L2 learner English stands out in that the frequency of *about* does not sit with the other written Englishes. It approaches but does not reach the spoken levels of the other corpora.

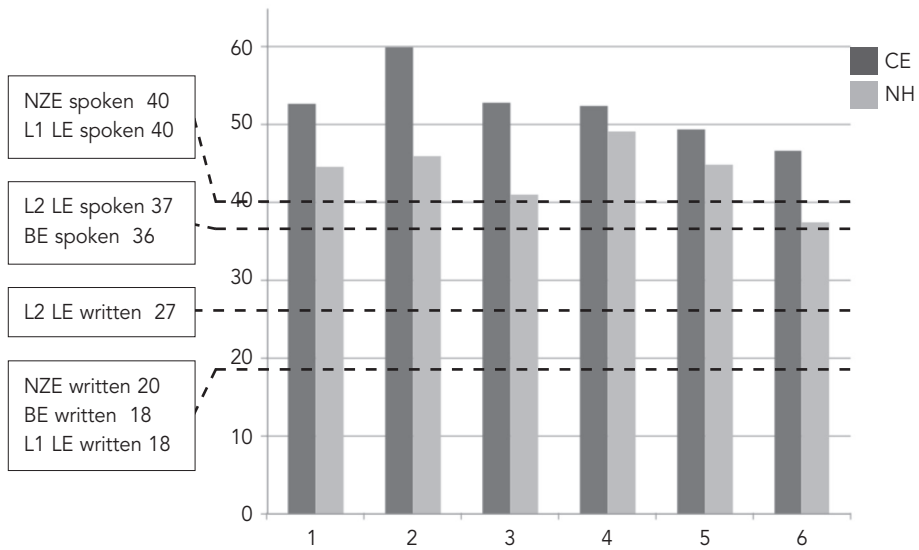When we consider uses of *about* in terms of the grammatical category of

NZE spoken 40
L1 LE spoken 40

L2 LE spoken 37
BE spoken 36

L2 LE written 27

NZE written 20
BE written 18
L1 LE written 18

Figure 1: *About* in CE and NH coursebooks and other corpora per 10k words

each use (see Table 9), we see that *about* occurs mostly as a preposition in all the corpora, and much less frequently as an adverb. The differences between preposition and adverb in each corpus are all statistically significant at the level of $p<0.0001$.

The following are examples of use in each grammatical category:

(1) Preposition: Write questions ***about*** the story using these question words. (NH series)

(2) Adverb: i think the minimum wage is ***about*** six dollars an hour or something that he can get (NZE spoken).

The two coursebook series use *about* as a preposition more frequently than L2 learners do, and both coursebooks and L2 learners use *about* more frequently as a preposition than BE, NZE and L1 learner English speakers and writers. There is no such clear pattern for uses of *about* as an adverb. Notice, for example, that the two series have rather different rates of use of *about* as an adverb.

Table 9: Part-of-speech in corpora per 10k words – *about*

	PREPOSITION	PREP *vs* ADV	ADVERB
BE spoken	27	χ^2 = 9466.328, df = 1, p<0.0001	8.7
BE written	14	χ^2 = 44546.643, df = 1, p<0.0001	4.1
NZE spoken	29	χ^2 = 767.206, df = 1, p<0.0001	12
NZE written	15	χ^2 = 457.508, df = 1, p<0.0001	5.4
L1 LE spoken	24	χ^2 = 25.138, df = 1, p<0.0001	15
L1 LE written	14	χ^2 = 185.470, df = 1, p<0.0001	3.8
L2 LE spoken	32	χ^2 = 1688.538, df = 1, p<0.0001	4.4
L2 LE written	24	χ^2 = 6182.201, df = 1, p<0.0001	2.9
NH series	39	χ^2 = 1108.864, df = 1, p<0.0001	4.8
CE series	41	χ^2 = 828.840, df = 1, p<0.0001	10

3.5.2 over

In the case of *over* (see Figure 2), the overall pattern noted earlier that there is no significant difference between the frequency of *over* in the two series, CE and NH (see Table 8) is evident also in the six books in the series. There is a trend in both series for the frequency of *over* to increase across the series, though that trend is not evenly spread across the series.

However, Figure 2 also shows that *over* is used less frequently in coursebooks than in all corpora except the L2 learner English written and spoken corpora. In other words, general English users and L1 learners use *over* more frequently than L2 learners, and coursebook English patterns are closer to the patterns for L2 learners than to those for L1 users of English both general and learner. This is unexpected in that we might expect coursebooks to use high frequency words even more frequently, if anything, than L1 users of the language, in order to make their texts more accessible to L2 learners.

When we consider uses of *over* in terms of the grammatical category of each use (see Table 10), we see that *over* occurs more frequently as a preposition than as an adverb in all the corpora except spoken NZE and spoken L2 learner English. The differences between preposition and adverb are all statistically significant at the level of p<0.0001, except for the CE and

NZE written 15
BE written 14
NZE spoken 13

BE spoken 10
L1 LE written 10
L2 LE written 10

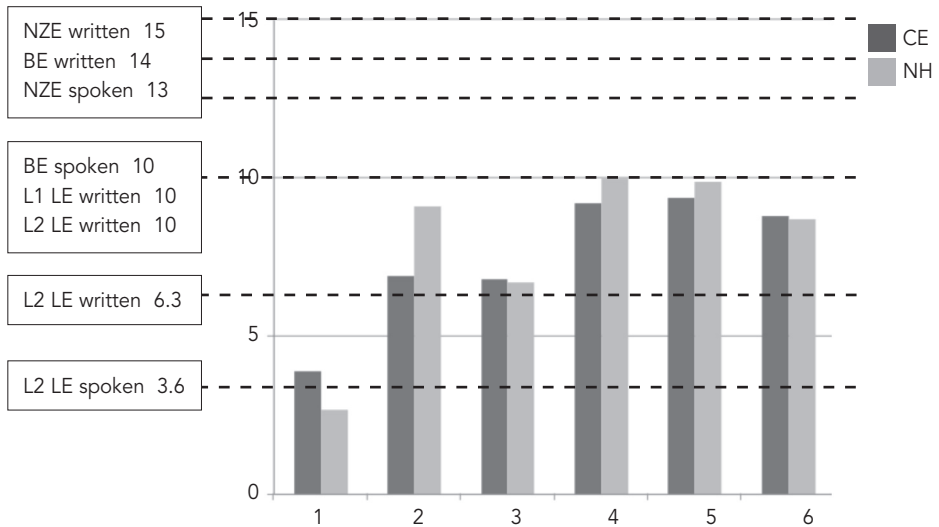L2 LE written 6.3

L2 LE spoken 3.6

Figure 2: *Over* in CE and NH coursebooks and other corpora per 10k words

NH coursebook series, where the differences are still significant but at lower levels of confidence. Uses of *over* as adjective and noun also occur in some of the corpora, but only at very low frequencies. The following are examples of use in each grammatical category:

(1) Preposition: You definitely shouldn't reveal the details of your bank account *over* the telephone. (CE series)

(2) Adverb: it was just when they all came *over* on the horses and it was just like oh wow this is a good film (L1 LE spoken)

(3) Adjective: oh God my university life is *over* (L2 LE spoken)

(4) Noun: At the start of the final *over*, England wanted five to win. (BE written)

The difference between preposition and adverb uses is much less marked for *over* than was the case for *about*; preposition uses of *over* are only slightly more frequent than adverb uses across the corpora. However, it is noteworthy that while preposition uses of *over* are consistently higher in the written corpora than the spoken, the reverse was the case for preposition uses of *about*.

Table 10: Part-of-speech in corpora per 10k words – *over*

	PREPOSITION	PREP *vs* ADV	ADVERB	ADJECTIVE	NOUN
BE spoken	5.6	χ^2 = 103.789, df=1, p<0.0001	4.6	0.03	<0.01
BE written	7.5	χ^2 = 1609.719, df=1, p<0.0001	5.9	0.02	0.02
NZE spoken	4.8	χ^2 = 76.849, df=1, p<0.0001	7.8	0.13	0.26
NZE written	9.3	χ^2 = 105.947, df=1, p<0.0001	5.3	0.61	0
L1 LE spoken	7.4	χ^2 = 23.902, df=1, p<0.0001	2.9	0	0
L1 LE written	8.7	χ^2 = 154.742, df=1, p<0.0001	1.7	0	0
L2 LE spoken	1.2	χ^2 = 25.610, df=1, p<0.0001	2.2	0.2	0
L2 LE written	3.8	χ^2 = 160.251, df=1, p<0.0001	2.2	0.03	0.02
NH series	4.6	χ^2 = 5.391, df=1, p= 0.0202	3.6	0.3	0
CE series	4.3	χ^2 = 6.260, df=1, p= 0.0123	3.3	0.2	0

3.5.3 end

In the case of *end* (see Figure 3), the overall pattern noted earlier that there is no significant difference between the frequency of *end* in the two series, CE and NH (see Table 8) is evident also in the six books in the series. There is a trend in both series for the frequency of *end* to increase across the series.

However, Figure 3 also shows that *end* is used less frequently in the first three coursebooks (but not NH Book 2) in both series and CE Book 4 than in all other corpora. In NH Book 2 occurrences of *end* reach the level of NZE written and spoken. In CE Book 5, occurrences of *end* reach the level of BE written, and in NH Book 4 they reach the level of L2 LE spoken and written. NH Books 5 and 6 and CE Book 6 reach the level of BE spoken and L1 LE written, but they are still well below the level of L1 LE spoken. These patterns suggest that coursebook usage builds towards L1 English usage.
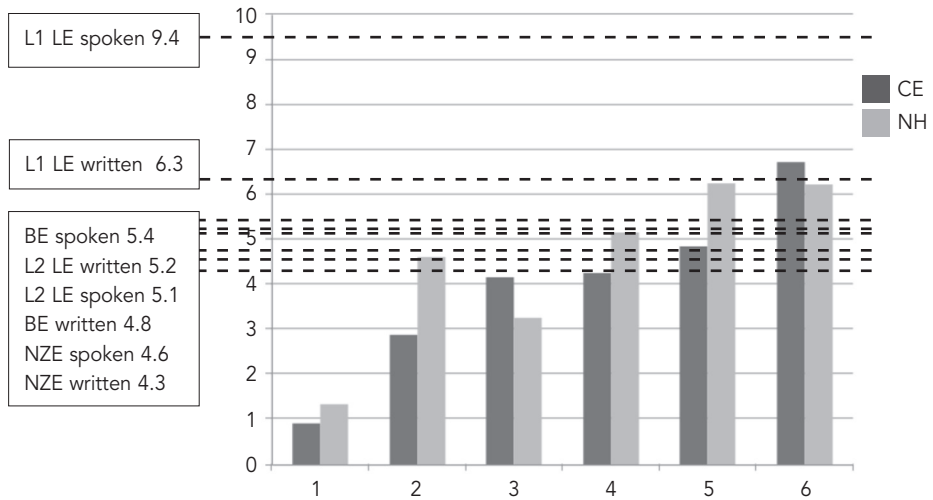
L1 LE spoken 9.4

L1 LE written 6.3

BE spoken 5.4
L2 LE written 5.2
L2 LE spoken 5.1
BE written 4.8
NZE spoken 4.6
NZE written 4.3

CE
NH

Figure 3: *End* in CE and NH coursebooks and other corpora per 10k word

Table 11: Part-of-speech in corpora per 10k words – *end*

	NOUN	NOUN *vs* VERB	VERB
BE spoken	4.7	$\chi^2 = 2985.196, df = 1, p<0.0001$	0.7
BE written	4.2	$\chi^2 = 24729.031, df = 1, p<0.0001$	0.5
NZE spoken	3.9	$\chi^2 = 217.815, df = 1,\ p<0.0001$	0.8
NZE written	3.9	$\chi^2 = 298.906, df = 1,\ p<0.0001$	0.3
L1 LE spoken	8.6	$\chi^2 = 77.919, df = 1,\ p<0.0001$	0.8
L1 LE written	4.8	$\chi^2 = 53.020, df = 1,\ p<0.0001$	1.6
L2 LE spoken	4.7	$\chi^2 = 271.488, df = 1,\ p<0.0001$	0.5
L2 LE written	3.7	$\chi^2 = 342.250, df = 1,\ p<0.0001$	1.5
NH series	3.2	$\chi^2 = 28.922, df = 1, p<0.0001$	1.4
CE series	3.3	$\chi^2 = 52.083, df = 1, p<0.0001$	1.0

When we consider uses of *end* in terms of the grammatical category of each use (see Table 11), we see that *end* occurs more frequently as a noun than as a verb in all the corpora. The differences between noun and verb are all statistically significant at the level of p<0.0001. The following are examples of use in each grammatical category:

(1) Noun: Just got to tie a bit of wood on the ***end*** of that bit where it's rotted. (BE spoken)

(2) Verb: The wood is split, and it isn't hard to ***end*** up with a splinter in your hand (NZE written)

Spoken L1 learner English stands out as having much more frequent occurrences of *end* as a noun than any of the other corpora, though this is not the case for verb uses. It may be the case that in this smaller corpus, for this word, topic has had an influence, in that the noun uses include many that relate to story plots, and an optional task in the interviews with learners was to talk about a film or play, and a required task was to describe four pictures making up a story (Gilquin et al. 2010: 8).

3.5.4 *place*

In the case of *place* (see Figure 4), the overall pattern noted earlier that there is no significant difference between the frequency of *place* in the two series, CE and NH (see Table 8), obscures some significant differences between books at particular levels. CE Book 1 has significantly (LL = 12.62, p<0.001) more occurrences of *place* than NH Book 1, as does CE Book 4 than NH Book 4 (LL = 5.26, p<0.05), while NH Book 6 has significantly (LL = 4.37, p<0.05) more occurrences of *place* than CE Book 6. There is a trend in the NH series for the frequency of *place* to increase across the series, while in the CE series the frequencies go up and down from book to book across the series.

However, Figure 4 also shows that, apart from NH Book 1, coursebook uses of *place* fall within much the same range as the L1 corpora, from BE spoken at 3.9 occurrences per 10,000 up to L1 learner English spoken at 6.8 occurrences. L2 learner English, both spoken at 6.4 and written at 8.0, fall towards or above the coursebook range. This is the only one of the four words which L2 learners have used more than L1 speakers of all sorts, and more than the coursebooks. It is possible that the learner corpus tasks/topics may have contributed to its higher use in the L2 corpora, but if that is the case, then it is odd that the same tasks in the learner written corpora have given different outcomes for L2 and
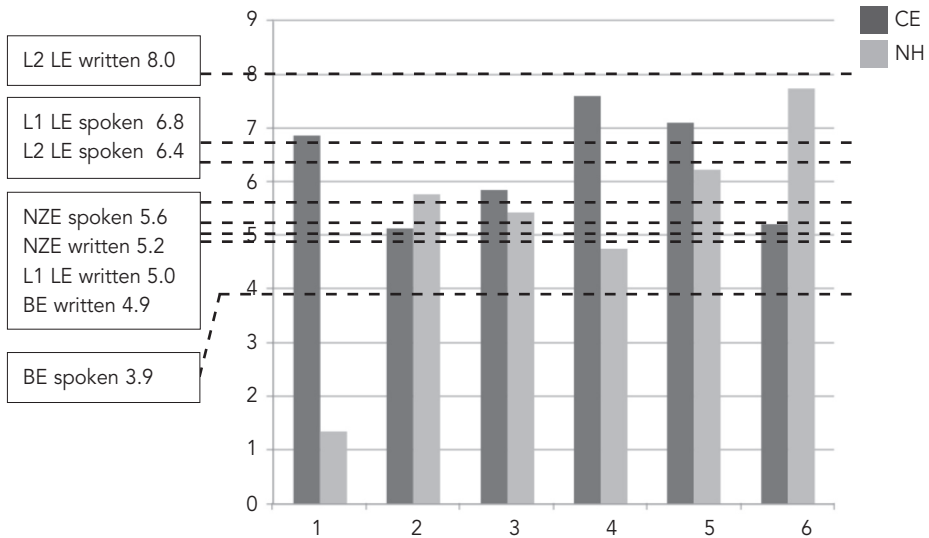
Figure 4: *Place* in CE and NH coursebooks and other corpora per 10k word

Table 12: Part-of-speech in corpora per 10k words – *place*

	NOUN	NOUN *vs* VERB	VERB
BE spoken	3.6	χ^2 = 2795.315, df = 1, p<0.0001	0.3
BE written	4.5	χ^2 = 27979.803, df = 1, p<0.0001	0.5
NZE spoken	5.4	χ^2 = 553.958, df = 1, p<0.0001	0.1
NZE written	4.9	χ^2 = 417.240, df = 1, p<0.0001	0.3
L1 LE spoken	6.8	χ^2 = 80.000, df = 1, p<0.0001	0
L1 LE written	4.7	χ^2 = 123.484, df = 1, p<0.0001	0.3
L2 LE spoken	6.3	χ^2 = 487.197, df = 1, p<0.0001	0.06
L2 LE written	8.0	χ^2 = 3019.000, df = 1, p<0.0001	0
NH series	5.5	χ^2 = 200.674, df = 1, p<0.0001	0.2
CE series	6.1	χ^2 = 247.918, df = 1, p<0.0001	0.2

L1 learners. The unexpectedly high comparative frequency of *place* may also have to do with its usefulness as a 'general noun' (Mahlberg 2005).

When we consider uses of *place* in terms of the grammatical category of each use (see Table 12), we see that *place* occurs more frequently as a noun than as a verb in all the corpora. The differences between noun and verb are all statistically significant at the level of $p<0.0001$. The following are examples of use in each grammatical category:

(1) Noun: It's also an extremely dangerous ***place***, and every year we have several cases where people, usually tourists have serious problems in the desert and have to be rescued. (CE series)

(2) Verb: Please ***place*** your tray tables in their fully upright and locked positions. (NH series)

There is no consistent pattern to the use of *place* as a noun in spoken and written corpora. BE and L2 LE show higher frequency in written than spoken use, while NZE and L1 LE show higher frequency in spoken than written use.

4. Pedagogical implications

We have seen that there is variation in the extent to which the frequencies of occurrence of different high frequency category-ambiguous words in general and learner English corpora are consistent with each other and with pedagogical applications in coursebooks. Such variation raises issues about common pedagogical practices.

Corpus frequency data should not necessarily be pedagogically prescriptive, but they should inform pedagogy. If coursebook writers and language teachers are well informed about issues raised by corpus comparison findings such as those presented here, they may choose not to be guided by them, but in that case they will do so in a principled way. For example, the findings of this study show that *about* is consistently used more in spoken than in written contexts, which suggests that pedagogical attention should follow that pattern. However, there may be a particular genre or topic in a written context that requires more attention to *about*. An example could be written fiction contexts where written dialogue resembles spoken usage.

Teachers need to consider what is going on where, for example, coursebook frequencies are lower than general English frequencies (e.g. *end*, particularly

as a noun, and *over*, particularly as a preposition) or higher (e.g. *about*, particularly as a preposition). Another example is where, for both 'content' (*end, place*) and 'function' (*about, over*) words, L1 learners are using these high frequency words more frequently than L2 learners in parallel corpora.

On the whole, it appears that the category ambiguity of these high frequency words is paid little explicit attention in the coursebook series. For example, in NH Book 2, *end* is included in a word list at the end of the book, and it is labeled as both noun and verb. However, there is no explicit attention paid to its category ambiguity in the relevant unit of work. It is simply the case that it occurs in both uses, noun and verb:

On 2 July she was nearly at the **end** of her journey, …

When did the Second World War begin / **end**?

Corpus comparison data suggest paying more pedagogical attention to the category ambiguity of high frequency words. The data also give language educators grounds for making reasoned choices about which category-ambiguous words and which uses of those words might warrant more, or less, pedagogical time and emphasis.

References

Audience Dialogue. 2006. Sustainable English - A vocabulary (Dennis List – 6 December 2006). http://www.audiencedialogue.net/susteng.html (last accessed on 20 June 2012).

Bauer, Laurie. 1993. *Manual of Information to Accompany The Wellington Corpus of Written New Zealand English*. Wellington, NZ: Department of Linguistics, Victoria University of Wellington.

Cunningham, Sarah and Peter Moor. 2005. *New Cutting Edge Intermediate Students' Book.* Harlow: Pearson Longman.

Cunningham, Sarah and Peter Moor. 2007. New *Cutting Edge Upper Intermediate Students' Book.* Harlow: Pearson Longman.

Cunningham, Sarah, Peter Moor and Jane Comyns-Carr. 2005. New *Cutting Edge Pre-Intermediate Students' Book.* Harlow: Pearson Longman.

Cunningham, Sarah, Peter Moor and Jane Comyns-Carr. 2007. *Cutting Edge Advanced Students' Book.* Harlow: Pearson Longman.

Cunningham, Sarah, Peter Moor and Frances Eales. 2005. *New Cutting Edge Elementary Students' Book.* Harlow: Pearson Longman.

Cunningham, Sarah, Chris Redston and Peter Moor. 2007. *Cutting Edge Starter Students' Book.* Harlow: Pearson Longman.

Davies, Mark. 2004. *BYU-BNC: The British National Corpus.* Available online at http://corpus.byu.edu/bnc

Gilquin, Gaëtanelle, Sylvie De Cock and & Sylvian Granger (eds). 2010. *Louvain International Database of Spoken English Interlanguage (LINDSEI)*. Louvain-la-Neuve: Presses Universitaires de Louvain.

Granger, Sylvian. 2009. 'The contribution of learner corpora to Second Language Acquisition and foreign language teaching: A critical evaluation'. In Karin Aijmer (ed.) *Corpora and Language Teaching*. Amsterdam: John Benjamins. 13–33.

Granger, Sylviane, Estelle Dagneaux, Fanny Meunier and Magali Paquot (eds). 2009. *International Corpus of Learner English, Version 2*. Louvain-la-Neuve: Presses universitaires de Louvain.

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

Kennedy, Graeme, D. 1987. Quantification and the use of English: A case study of one aspect of the learner's task. *Applied Linguistics* 8, 264–286.

Leung, Matthew Wing Kwong. 1991. Prototype theory and teaching English prepositions. *Perspectives (City University of Hong Kong)* 3, 89–97.

Mahlberg, Michaela. 2005. *English General Nouns. A Corpus Theoretical Approach*. Amsterdam: John Benjamins.

Meunier, Fanny and Céline Gouverneur. 2009. New types of corpora for new educational challenges: Collecting, annotating and exploiting a corpus of textbook material. In K. Aijmer (ed.) *Corpora and Language Teaching*. Amsterdam: John Benjamins, 179–201.

O'Keeffe, Anne, Michael McCarthy and Ronald Carter. 2007. *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press.

Scott, Mike. 2004. *WordSmith Tools version 4*. Oxford: Oxford University Press.

Soars, Liz and John Soars. 2002. *New Headway Beginner Student's Book*. Oxford: Oxford University Press.

Soars, Liz and John Soars. 2003a. *New Headway Advanced Student's Book*. Oxford: Oxford University Press.

Soars, Liz and John Soars. 2003b. *New Headway Intermediate Student's Book (3rd ed.)*. Oxford: Oxford University Press.

Soars, Liz and John Soars. 2005. *New Headway Upper-Intermediate Student's Book (3rd ed.)*. Oxford: Oxford University Press.

Soars, Liz, John Soars and Sylvia Wheeldon. 2000. *New Headway Elementary Student's Book*. Oxford: Oxford University Press.

Soars, Liz, John Soars and Sylvia Wheeldon. 2007. *New Headway Pre-Intermediate Student's Book (3rd ed.)*. Oxford: Oxford University Press.

REVIEW

Barbour, Julie. 2012. *A Grammar of Neverver.*
Berlin: De Gruyter Mouton.

Elizabeth Pearce: *School of Linguistics and Applied Language Studies,
Victoria University of Wellington <elizabeth.pearce@vuw.ac.nz>*

Approximately 40 languages have been identified as distinct indigenous languages of the island of Malakula, Vanuatu (Lynch and Crowley 2001: 4). With Julie Barbour's description of the Neverver language there is now a total of nine Malakula languages for which a reasonably substantial published description is available.[1] It is a very substantial treatment (covering 476 pages) and the work brings a most welcome contribution to knowledge of Malakula languages and of Vanuatu languages more generally.

The Neverver language description is presented in 13 chapters. It has an index and five appendices, the latter including an interesting account of correspondences between the contemporary data and earlier material collected by Bernard Deacon (Deacon 1926–1927), an assessment of the vitality of the language and four sample texts. My discussion of the content of the chapters of the book that follows will be rather selective in that my focus will be on bringing out details in the description which I find to be of particular interest in the context of phenomena that have or have not been observed in other languages of Vanuatu, or, more especially in languages of Malakula.

Chapter 1 sets the scene with an informative discussion of the situation of the Neverver language, its location and its speakers, along with a review of previous research on the language.

Chapter 2 covers the phonetic characteristics of the segment inventory and the phonological processes that are implicated in the analysis. Although Neverver lacks labio-velars, it has an interesting array of consonant phonemes. Neverver has both /l/ and /r/ and a number of prenasalized consonant segments: the usual voiced stops, a bilabial trill, a voiced alveo-palatal affricate and an alveolar trill (the latter distinct from the plain trill). Geminate phonemes are in contrast with non-geminates for the liquids, nasals and voiceless plosives. The phonemic contrasts are exemplified in (near) minimal pairs and we are shown acoustic measurements for length supporting the geminate/non-geminate contrasts. We also see interesting evidence justifying the distinctions between segment sequences versus single segment complex phonemes.

One of the phonemic contrasts that might possibly have been more robustly detailed, however, is that between prenasalized /dʒ/ and the prenasalized trill /D/. In the presentation of these segments, the (near) minimal pairs show /dʒ/ in contrast with /s/, but /D/ is shown contrasting with /d/ and /r/. A distinguishing characteristic of /dʒ/ is that it occurs as [ⁿs] in final position whereas a variant of /D/ in final position is [r]. In the data that are shown on these final position occurrences, however, the contexts are not exactly parallel: vowels preceding /dʒ/ are /i/ and the diphthong /au/ (p. 36), whereas final /D/ is shown preceded by /a/, /e/ and /o/ (p. 38). In syllable-initial position, we see one occurrence each of both [d] and [D] followed by /o/, but the remaining data shows [dʒ] followed by /a/ and /e/ and [D] by /i/ and /u/. This means that it is only in the /#__ o case that we have what counts as the most plausible evidence for the claimed contrast between /dʒ/ and /D/. It would have been useful to see explicit comparison of these two segments if the relevant forms for such comparison are attested in the language.

Also of interest are a number of further phonological processes. Neverver has instances of metathesis and it has both [i] epenthesis (with fast speech realizations as [ə] in a set of restricted environments) and consonant excrescence. There is final devoicing and final /p/ and /k/ are realized as their fricative counterparts. Before /u/, /p/ is realized as a voiceless bilabial trill. The demonstrations of phonotactic effects and other phonological processes are clearly set out in autosegmental CV tier diagrams.

There are cases in the analysis of phonological data in languages where there is indeterminacy around what is the best choice for the representation of the underlying segment. For Neverver, a particular case of this type concerns the representation of the Irrealis prefix which surfaces in different contexts as [m] or as [ᵐb]. The question is whether this prefix is best analyzed as /m/ with

consonant excrescence in certain environments, or as /^{m}b/ with weakening to [m] in other environments. The solution that is argued for takes the consonant excrescence approach, with the rule statement (p. 64):

(1)   m: [^{m}b] / __ r, l, j, V

In support of this analysis is the realization of epenthetic plosives, including [d] and [g] in contexts where a morpheme-final nasal is followed by a vowel. Running somewhat against this analysis is the different behaviour of the stative prefix /m/ which induces consonant excrescence before the approximant and variably before the liquids, but not before vowels (p. 64). Further discussion of the Irrealis prefix is given in section 6.1 where the variant realizations of this prefix are set out in detail. Here we see the data on alternations in contexts with following plosives in which the Irrealis prefix is realized simply as [m] or, where intervening vowel epenthesis induces the realization, as [^{m}b]. An obvious alternative treatment of the data (and one which I find more plausible) would take the Irrealis prefix to be /^{m}b/ with weakening to [m] in the absence of vowel epenthesis when the following consonant is a plosive. Lynch (2007) has proposed that comparative Malakula data points to the Irrealis markers as probably being reconstructable as *$b^{(w)}a$. Whether Neverver presents as an instance of historical weakening to /m/ for its Irrealis prefix, or whether this prefix is still a plosive in the grammars of its speakers is an interesting question in the consideration of the processes involved in the trade-off between the interacting roles of epenthesis and weakening/strengthening processes.

Chapters 3–5 present the description of nouns and pronouns and other components of noun phrases. The Neverver pronoun system maintains the Proto-Oceanic inclusive/exclusive contrast but, whilst the subject-agreement marking on the verb maintains the three-way singular/dual/plural distinction, the pronoun paradigm encodes only the singular/non-singular contrast. Nouns are divided into three classes: common nouns, personal nouns and local nouns (p. 76). Unlike in some other Malakula languages (Pearce 2007; Lynch 2007), the incorporation of the Proto-Oceanic article *na has applied to common nouns in Neverver across-the-board, giving the result that, for the most part, aside from borrowings, such nouns in Neverver have nV onsets.

Very unusual for a Vanuatu language are the parameters that define the distinctions in form in the Neverver possession marking system. Nouns with synchronic final n are immediately followed by the possessor phrase. The form of the linker between other nouns and a possessor phrase is a function of whether the possessor is human or not. Human possessor phrases are preceded

by the genitive linker *titi* and, in possessor phrases with pronoun heads, the pronouns have a *t-* prefix (with *titi* as the third person singular form). Nonhuman possessor phrases are preceded by *an*, which is homophonous with the relative clause complementizer. Whilst the phonologically conditioned distribution of forms of possession cuts across the human/nonhuman divide, the alienable/inalienable distinction is manifested in that it is most often the case that nouns with final *n* can be viewed as nouns which inherently are inalienable (body parts, etc.). The final *n* appears to be a reflex of an earlier third person singular possessive suffix. Such singular suffixes are otherwise now relics (understood but not produced by younger speakers): as *-g* '1SG', *-m* '2SG' and *-n* '3SG' (p. 134). The synchronic outcomes of the historical third person singular *-n* incorporation now result in mismatches in forms of the type: *nevran na* 'hand.3SG 1SG' = 'my hand' (p. 132).

Adjective members of a restricted set immediately follow the noun, but other modifier expressions (including numerals) occur in relative clauses or reduced relative clauses. Full relative clauses with definite reference are introduced by the complementizer *an* and with *ang~a* 'anaphoric' at their right edge (the right edge marker can be omitted when the referent is a nonsubject). Only the numerals ≤ 9 bear the third person singular verbal agreement prefix, whether occurring in isolation or as parts of complex numbers (e.g. *nangavul **i**-ru nidruman **i**-skham* 'ten 3SG-two plus 3SG-one' = '21'). The detailing of noun phrase-internal ordering shows that relative clauses occur finally followed only by the delimiter *ang/a* and that possessives take up a position following lexical modifiers and intensifiers and before quantifiers, demonstratives and the plural marker in that order. Possessor arguments can also be preposed to come before the head noun.

Chapters 6–8 describe the verb morphology and verb complexes showing how the different forms express a range of agreement and tense/aspect and mood functions.

As has already mentioned, verbs standardly encode the inclusive/exclusive distinction and the three-way number distinction singular/dual/plural in their subject marking. There is a further prefix encoding impersonal subjects which has distinct forms in the Realis and the Irrealis. There are cases of stem-initial mutated forms, one category being phonologically conditioned and one category exhibiting a transitive/intransitive alternation. Verbs can bear an applicative suffix, but monotransitive verbs are otherwise unsuffixed. Whilst Neverver lacks a syntactic passive construction, it has a lexical passive verb *bal* 'be hit' (p. 186).

Chapter 7 presents the description and classification of various markers encoding tense, mood and aspect on and around the verb. The discussion in this chapter is especially well supported with the inclusion of input from a range of sources treating the understanding of these kinds of functions from the cross-linguistic perspective. Whilst Realis/Irrealis mood is encoded in the verb prefixes, aspectual and tense denoting functions are manifested as verb suffixes or as postverbal particles. Overall, across these two categories, the language has a total of eleven distinct such markers. The exposition covers some rather fine-grained distinctions for these functions and is very well carried through.

Chapter 8 is devoted to reduplication in verb forms from both formal (phonological) and functional perspectives. The chapter provides a well detailed treatment of the topic.

Chapter 9 lays out the basic patterns of clause structure, including coverage of negation, interrogatives, reflexives and reciprocals, impersonal constructions, comparatives, clause modifiers and non-verbal propositions. The Neverver clause has a basic SVO ordering. Subjects are not required to be overt and objects can also be omitted. There are two ditransitive constructions in which the Recipient argument precedes the Theme argument: (i) where the verb occurs as a compound -*lav-lik* 'get-pass'; and (ii) where the verb takes the applicative suffix -*ikh*. Interrogative constituents may occur in-situ or they may be preposed. The different categories of propositions which can be expressed in clauses lacking a predicate marked by subject-agreement are all matched by counterpart constructions employing a lexical verb.

Chapters 10 and 11 discuss a range of constructions that can be characterized as varieties of Serial Verb constructions. The division in the presentation between constructions with Complex Nuclei (Chapter 10) versus those with Complex Cores (Chapter 11) follows the classification set out in Foley and Olsen (1985). In essence, the Complex Nuclei constructions have two contiguous predicates, whereas in constructions with Complex Cores the two predicates are not required to be contiguous and the second predicate bears a subject-agreement prefix. In both cases, the constructions are considered to be mono-clausal on a number of tests (there is, for instance, no possible distinction in mood marking on the two predicates). The Complex Core constructions can be further subdivided in accordance with the available subject-agreement marking on the V2 of the construction: either (default) third person marking, obligatory coreferentiality between the V1 and V2 subject, or availability of non-coreferential subject-agreement marking on V1 and

V2. The different construction types are carefully detailed and the discussion provides an interesting contribution of data and analysis for the study of complex predicate constructions more generally.

In Chapter 12, complement clause constructions are classified in terms of the semantics of the higher predicates. The complement clauses may or not be preceded by one of two available complementizers (*il* or *at*), with certain predicates showing preferences for one over the other and with some possible meaning distinctions in the presence/absence of a complementizer in one or two cases. An interesting aspect of the discussion in this chapter concerns the polarity effects where a negated higher predicate imposes Irrealis marking on the verb of its complement clause.

Further kinds of complex sentences including more than one clause are described in Chapter 13: adverbial clauses and constructions with conjoined clauses. Certain intonation patterns are relevant for some distinctions in meaning and/or discourse function. Barbour points out (p, 411) that the basic conjunctive coordinator of the form *ga* has phonological identity with segments of *gang* 'be like so', and with *baga* 'then', and occurs also in the reduplicated form *gaga* 'on and on until'. The adversative linker *be* is a likely borrowing from Bislama. To this category I would also add the linker *we* which is shown as being used before a repeated verb in an augmentative/ intensifying function (pp. 414–415), comparable to one of the uses of Bislama *we* as documented in Crowley (2003).

For a grammar of a language, quite a bit of space in the book is taken up with input from a variety of sources spelling out the bases for the terms and classifications that are adopted. One has a bit the impression of gleanings from a range of offerings in the typological literature. This type of input, however, as well as providing the underpinnings of the descriptive apparatus, has the useful function of locating the phenomena that is documented for Neverver in terms of its cross-linguistic comparability. At a number of points in the description, we are also given insights into how Neverver stacks up against other Vanuatu/Malakula languages with respect to comparable morphology or syntax. To a lesser extent, particular aspects of Neverver morphology are represented with respect to hypothesized Proto-Oceanic sources. A very commendable feature of the book is the extensive input about intonation patterns which are relevant to defining and distinguishing phrases of various types. The book scores highly on readability. The explanations about the classifications and the analyses of data are presented with great clarity.

The aim of the series in which the book appears, the Mouton Grammar

Library, is stated on the back cover as that of '. . . build[ing] an extensive collection of high quality descriptions of languages around the world'. There is no doubt at all that Barbour's description of Neverver is well consonant with the definition of being a high quality description. It is a substantial and most welcome addition to the documentation of languages of Vanuatu and it makes an especially informative contribution in filling one of the gaps in our knowledge of the languages of Malakula.

Notes

1   Lynch and Crowley (2001: 17-19) cite just two Malakula languages (Port Sandwich and V'ĕnen Taut) as having been reasonably well described. The following six languages can now also be included in this category: Naman (Crowley 2006a), Avava (Crowley 2006b), Neve'ei (Musgrave 2007), Tirax (Brotchie 2009), Navahaq (Dimock 2009) and Unua (Pearce 2015). Available now also are descriptions of a lesser substance of Tape (Crowley 2006c) and of Nese (Crowley 2007d) and a sketch grammar of Ninde (Dimock et al 2014).

References

Brotchie, Amanda. 2009. *Tirax Grammar and Narrative: An Oceanic Language Spoken on Malakula, North Central Vanuatu*. PhD thesis, University of Melbourne.

Crowley, Terry. 2003. *A New Bislama Dictionary.* Second edition. Vila: Pacific Languages Unit, University of the South Pacific.

Crowley, Terry. 2006a. *Naman: A Vanishing Language of Malakula (Vanuatu)*. Unpublished manuscript, University of Waikato.

Crowley, Terry. 2006b. *The Avava Language of Central Malakula (Vanuatu)*. [ed. by John Lynch]. Canberra: Pacific Linguistics.

Crowley, Terry. 2006c. *Tape: A Declining Language of Malakula (Vanuatu)*. [ed. by John Lynch]. Canberra: Pacific Linguistics.

Crowley, Terry. 2006d. *Nese: A Diminishing Speech Variety of Northwest Malakula (Vanuatu)*. [ed. by John Lynch]. Canberra: Pacific Linguistics.

Deacon, A. Bernard. 1926–1927. *Fieldnotes on the Nesan Language*. The Haddon Papers, 16,008–16,015, 17,000. Cambridge: Cambridge University Library.

Dimock, Laura Gail. 2009. *A Grammar of Nahavaq (Malakula, Vanuatu)*. PhD thesis, Victoria University of Wellington.

Dimock, Laura, Emily Greenbank, Leina Isno and Elizabeth Pearce. 2014. *A Sketch Grammar of Ninde*. http://catalog.paradisec.org.au/collections/EP1.

Foley, William and Mike Olsen. 1985. 'Clausehood and verb serialization'. In Johanna Nichols and Anthony C. Woodbury (eds.) *Grammar Inside and Outside*

the Clause: Some Approaches to Theory from the Field. Cambridge: Cambridge University Press. 17–60.

Lynch, John. 2007. 'The Malakula linkage of Central Vanuatu'. Paper presented at the *Seventh International Conference on Oceanic Linguistics (COOL7).* Noumea, 2–6 July.

Lynch, John and Terry Crowley. 2001. *Languages of Vanuatu: A New Survey and Bibliography*. Canberra: Pacific Linguistics.

Musgrave, Jill. 2007. *A Grammar of Neve'ei, Vanuatu*. Canberra: Pacific Linguistics.

Pearce, Elizabeth. 2007. The reflexes of Proto-Oceanic **na* in Unua. In Jeff Siegel, John Lynch and Diana Eades (eds.) *Language Description, History and Development: Linguistic Indulgence in Memory of Terry Crowley.* Amsterdam: Benjamins. 327–339.

Pearce, Elizabeth. 2015. *A Grammar of Unua*. Berlin: De Gruyter Mouton.

TEACHING LINGUISTICS AND SOCIOLINGUISTICS IN NEW ZEALAND

BRINGING LINGUISTICS TO LIFE: AN ANCHORED APPROACH TO TEACHING LINGUISTICS TO NON-LINGUISTS

Arianna Berardi-Wiltshire: *Linguistics and Applied Linguistics, Massey University <a.berardi-wiltshire@massey.ac.nz>*

Peter Petrucci: *Linguistics, Massey University <p.r.petrucci@massey.ac.nz>*

Abstract

Massey University has a long established programme in Linguistics in the Bachelor of Arts, in both face-to-face and distance modes. With the recent introduction of the Bachelor of Communication, linguistics-teaching staff find themselves working with a new and growing cohort of students whose specific needs and interests unambiguously straddle the arts/social sciences *and* business. This article examines two courses that form part of the BC major and minor in Linguistics. Activities from these papers are presented as examples of the ways in which our course coordinators endeavour to meet the pedagogical challenges of 'teaching linguistics to non-linguists' by applying the principles of Anchored Instruction (Bransford et al., 1990) to embed linguistics in current issues and contexts that are authentic and often familiar to students, and by utilising students' existing skills and intuitions about language as a valuable starting point for learning and as a key to unlocking a positive sense of belonging to an established linguistics community of practice (Lave & Wenger, 1990). We conclude with a reflection on the need to adapt teaching philosophies and practices to the ever-evolving needs of the student population by continually re-packaging linguistics in a new light without compromising the integrity of the discipline, an endeavour that we feel has profound and on-going implications for our identities as linguistics educators.

1. Introduction

Since its beginnings in the now remote mid-1970s, the distinctive identity of the Linguistics programme at Massey University has been largely shaped by two main factors, which throughout the years have influenced not only the range of papers offered and their content, but also, and most pertinently, the pedagogical philosophies and practises of the teaching staff. A first, historical factor is to be found in the nature of Massey University as a provider of distance education as well as of traditional face-to-face courses. A second, more recent influence is the inclusion of Linguistics as a major in the recently established Bachelor of Communication (BC), a joint degree shared by the College of Business and the College of Humanities and Social Sciences. Throughout the evolution of the programme to the present day, we have endeavoured to meet the pedagogical challenges of 'teaching linguistics to non-linguists' in a context whereby the inherent complexity of the task is compounded by the need to do so across internal and distance mode, achieving the mandated degree of cross-mode equivalence and without compromising the integrity of the discipline.

   This paper discusses our approach to teaching Linguistics by examining learning activities from two offerings 172.236 Forensic Linguistics and 172.232 Language and Society in New Zealand, two papers recently tailored to align more closely with the interests of our BC students. Activities from these papers are presented here to illustrate the ways in which our course coordinators successfully apply the principles of *Anchored Instruction* (Bransford et al., 1990) to embed linguistics in current issues and contexts that are authentic and often familiar to students, and by utilising students' existing skills and intuitions about language as a starting point for learning and as a key to unlocking a positive sense of belonging to an established linguistics community of practice (Lave & Wenger, 1990). The ultimate aim of the paper is to offer our experience with the principles of Anchored Instruction in teaching linguistics within our specific setting as a way to illustrate how such principles can be employed in the linguistics classroom to overcome some of the challenges of teaching linguistics to 'non-linguists', or in other words, to students who are not pursuing linguistics qualifications, and/or that have little or no previous knowledge of the subject.

   The concluding comments offer a reflection on the necessity to adapt teaching perspectives and practices to satisfy the ever-evolving needs of a

complex student population Ultimately, in our experience, this necessity requires teachers to continually re-package linguistics in a new light without compromising the integrity of the discipline, an endeavour that we believe has profound and on-going implications for our identities as linguistics educators.

2. Linguistics in the BA and BC

The beginnings of Linguistics as a subject in the Bachelor of Arts (BA) saw the discipline taught through a traditional structured programme including a range of courses focused around the discipline's main subfields, i.e. phonetics, syntax, sociolinguistics, etc. In the early 2000s, in response to changes and fluctuations in the student population associated with a marked increase of international students, attempts were made to restructure the linguistics programme so to appeal to a wider range of students both within the humanities and the social sciences, whose students had the option of taking linguistics papers as electives. Primarily, this involved efforts towards framing the instruction of traditional linguistic theories and issues in ways that made them both relevant and interesting not only to students wanting to major in linguistics or traditionally related subjects such as English or European languages, but to potentially all students in the BA, irrespectively of their chosen major. Naturally this also involved a rethinking of both content and instruction strategies that would address the need to cater for internal students as well as large cohorts of distance students, maintaining a strict degree of teaching equivalence across the two modes. Roughly at the same time, the university wide adoption of new distance teaching technologies contributed to the need for innovation, offering new opportunities, but also requiring from teachers consistently high levels of flexibility and adaptability to unfamiliar strategies and tools.

Then, more recently in 2005, Linguistics as a subject was included in Massey's new Bachelor of Communication (BC), a joint degree offered through the colleges of Business and Humanities and Social Sciences. The university website describes the BC as a degree that combines Business with Humanities 'to create a well-rounded innovated communication specialist who is sought after by employers' ('BC', n.d.). Students who pursue this degree need to major in one subject from one college and minor in another subject in the other college. Possible majors from the College of Business are Communication Management, Journalism Studies, Marketing Communication

and Public Relations. From the College of Humanities and Social Sciences, possible majors include Expressive Arts, Media Studies, and Linguistics. So, for instance, a BC student might major in Linguistics and minor in Journalism Studies, or the reverse, major in Journalism Studies and minor in Linguistics.

With the inclusion of Linguistics in the BC, the programme saw the emergence of a very specific and complex set of teaching requirements arising from the combination of teaching linguistics to both majoring and non-majoring students of both BA and BC students in both internal and distance mode. Particularly pressing in this regard are the challenges associated with a current student population consisting of:

1. Students who, for the most part, have never even heard of linguistics as a subject, and have never envisaged linguistics as a potential study subject and/or career path. For these students, explanations of linguistics need to be accessible and interesting, possibly relating to areas of their personal experience.

2. Students who did not set out to major in linguistics, but that are instead pursuing majors in other subjects such as Psychology, Developmental Studies, Journalism, Media Studies and Communication Management. To these students, linguistics is made particularly interesting and worthwhile when introduced in its conceptual associations and applications to fields of knowledge they are pursuing, which most often lie beyond the humanities and social sciences.

3. Students who fall into the 'non-traditional' category, pursuing either a BA or BC through distance mode. These students tend to be more mature, both in terms of age and life experiences, and many are already in well-established careers in a wide range of contexts. Over the years this group has included, for example, police and customs officers, government consultants and language teachers in New Zealand and abroad. While these students tend to be more readily able to see the applications of linguistics outside purely academic contexts, they require specific and consistent efforts to build and maintain interest, motivation and engagement through teaching strategies that offset the sense of social isolation and lack of personal support that can be common to distance learners.

3. The Anchored Instruction approach

Anchored instruction is a technology-based type of situated-learning approach originally developed by John Bransford and The Cognitive and Technology Group at Vanderbilt University in the US state of Tennessee (Branford et al., 1990), which stresses the importance of placing learning within a meaningful, problem-solving context. Within anchored instruction, an *anchor* is a story, a context or a specific situation that includes a problem or issue to be resolved and that is designed to be of interest to the learners. The 'anchoring' reference has to do with the grounding of the teaching content through a realistic and authentic scenario, making learning meaningful for students, often allowing them to experience the same dilemmas facing experts in a given field.

In typical anchored instruction classrooms, anchors are presented through a brief video clip, and the students work together in small groups to formulate strategies for solving the problem embedded in the anchor. During this collaborative stage of the activity, students are allowed to take ownership of the problem and to become actively involved in generating a solution, while the teacher's role shifts to that of a facilitator, helping the students through the extraction and organisation of data related to the problem, allowing the students to struggle, but supporting the process when necessary. Later, on completion of the activity, the teacher might use analogous scenarios or problems to help students understand issues more deeply by exploring the relationship among different variable (e.g. *What if you encountered this same situation but instead of X you had Y?*). Similarly, extension problems requiring skills or strategies similar to those used in the initial scenarios can also be used to expand and/or strengthen learning.

Similar to problem-based learning and case-based learning, anchored instruction is based on cognitive constructivist perspectives that stress the importance of 'situated cognition' (Brown, Collins, & Duguid, 1989) and 'cognitive apprenticeship' (Collins, Brown, & Newman, 1989). Specifically, the approach was proposed as a way to engage students in activities that can help reduce the 'inert knowledge' phenomenon, whereby ideas 'are merely received into the mind without being utilised, or tested, or thrown into fresh combinations' (Whitehead, 1929:1). Overall this principle suits the teaching of linguistics well, and indeed many areas of formal linguistics are commonly delivered largely through problem-based or dataset-based lessons in traditional linguistics programmes. The point of difference with Anchored

Instruction is that the way the problem-solving task is designed, packaged and introduced to the students using anchors that encompass situations, stories, characters that are purposely designed to appeal to the students, presented though familiar media. Within such anchors, linguistic knowledge and skills are contextualised to issues and problems that are likely to be relevant to the students' own life and experiences, making it easier for them to internalise relevant concepts and to recognise when to appropriately apply these to solve similar problems in the future.

As with other forms of situated learning, anchored instruction is based on a social constructivist perspective whereby learning is seen as a function of the activity, context and culture in which it occurs, and where social interaction is crucial to learners' access to knowledge and behaviours that are specific to a particular community of practice (Lave & Wenger, 1991). As a beginner moves from the periphery of this community to its centre through activities of an interactive and collaborative nature, they become active and engaged with the community's culture, gradually gaining the expertise associated with full membership. Of course in order for this to happen, novices must be presented with discipline-specific content that approximates and/or reproduces the ways in which knowledge is used in real-world applications by members of the community, as it is only through authentic learning environments that learners can develop the knowledge and skills necessary to 'create, innovate and communicate' (Herrington & Herrington, 2007:69) like an expert in the field.

The principles of anchored instruction are easily applied to learning activities designed to teach linguistics both in the classroom and within online environments aimed at distance learning, where linguistics concepts, theories and applications can be introduced and illustrated by anchors that are readily available via YouTube and other online sources, and with which the students are often already familiar. On the first day of Language and Communication, for instance, students are made to engage in a preliminary analysis of the language used in a current and well-known drink driving commercial targeted at young viewers. The conclusions they come to, with some guidance from the instructor but without the use of any previously introduced linguistic concept or analytical tool, serve as a launching point into a discussion of linguistics and communication.

Even in this illustrative rather than strictly problem-solving form, our experience with elements of anchored instruction suggests that this is an effective approach that can help address many of the challenges of teaching linguistics to non linguists across different modes. Specifically, in our

experience, the use of an anchored instruction approach in our linguistics classes have helped with:

1. making information and learning more relevant, useful, and personally meaningful to our students, introducing linguistic concepts in an immediate and accessible way that can stimulate interest and facilitate learning, empowering learners to engage with the discipline independently from their academic background and level;

2. making linguistics interesting and worthwhile by introducing it in its conceptual associations and applications to the other fields of knowledge and/or experience, giving students an appreciation for the power of common concepts and cross contextual applications;

3. encouraging the adoption of multiple perspectives and social collaboration in problem-solving activities that highlight the value of individual viewpoints, promote peer-to-peer connections and contribute to stimulating and supportive learning environments, both in the classroom and by distance.

The following sections will illustrate the above points with specific examples of how the principles of anchored instructions are currently used to teach linguistics in our internal and distance courses within the BA and the BC.

4. Anchored instruction in practice

Within the programme, face-to-face teaching tends to follow a traditional lecture/tutorial/assessment format. For distance students, written study guides replace internal lectures and online forums act as the main means of student-to-student and student-to-teacher interaction. *Moodle*, Massey's online learning interface used for distance teaching, allows for the sharing of multimedia content, and increasingly distance students are also being given access to recorded lectures and to synchronous online tutorials using interactive technology such as Scopia or Adobe Connect.

   Whether delivered in a classroom or online, the content of our linguistics courses is for the most part immediately linked to specific contexts and real-life situations through learning activities designed according to the principles of anchored instruction discussed above. For instance, in addition to the drink-

driving commercial mentioned above, the Language and Communication paper introduces language functions and speech acts by guiding the students through an analysis of the linguistic features in Air New Zealand inflight safety videos available on YouTube. Similarly, in Language and Society in New Zealand, students are introduced to the key factors in language maintenance and shift through a problem-based activity designed around a series of video interviews with New Zealand migrants discussing their efforts to maintain their mother tongues, also available on YouTube.

The anchors used in these activities diverge from those in traditional anchored instruction classrooms in that they are not tailor-made for the specific activities they are to be a part of; however, we believe that this does not compromise their suitability in any way. As a matter of fact, in our discipline, where the authenticity of linguistic data is so often crucial, anchors of a non-customised nature can be just as—if not more—effective than tailor-made ones in supporting students' learning by effectively demonstrating the relevance of linguistics to all areas of life where language is involved.

The basic principles of anchored instruction are so adaptable that they can be used not only to introduce new linguistic concepts in lectures, but also for extension activities in tutorials and even in tasks designed to assess different types of discipline-related proficiencies. To illustrate this point, the following two sections will examine an example of anchored instruction in a tutorial activity and as part of the preliminary stages of a course assignment task.

4.1 Forensic Linguistics tutorial activity: Acoustic analysis and authorship

Forensic Linguistics is a second-year paper popular with BC, BA and students of many other disciplines, including sciences and human health subjects. Because the prerequisite is any first-year BA paper, rather than an introductory linguistics paper, some students who enrol in Forensic Linguistics have not studied linguistics before. Consequently, the paper also acts as an introduction to linguistics. The course is broken down into four main topics, Trademark Issues, Product Liability, Questioned Authorship, and the Use/abuse of Language in Legal Contexts, all concurrently presented with information about particular linguistic subfields. For instance, the unit exploring Questioned Authorship includes segments on acoustic phonetics, syntax, corpus linguistics and pragmatics. The discussion of linguistic concepts for each unit is not exhaustive. Instead, much the way a forensic linguist acting as an expert witness might introduce linguistic concepts to

a jury (McMenamin, 2002), the linguistic material included in the paper is that which is necessary for understanding the topics and actual forensic cases under discussion.

The tutorials are designed to allow students to apply forensic analytical procedures to real-life contexts. Many of the situations that act as anchors in the tutorials are freely available on YouTube and other video-sharing platforms. One such tutorial, covering spoken authorship and intended meaning, is designed around a controversy involving an American politician on the 2012 presidential campaign trail. At a campaign event in the state of Iowa, the politician addressed welfare reform and allegedly made the following racially insensitive comment in front of the national media: 'I don't want to make black people's lives better by giving them somebody else's money.' He later denied making this comment and pointed out that welfare recipients were not limited to a specific ethnic group. He claimed that what he had actually said was not *black*, but rather *blah*, to signal a change in his line of thinking for the utterance at hand. The controversy appeared on YouTube in various forms shortly after it happened.

Figure 1: *Rick Santorum said blah or black* (YouTube, n.d.)

In the tutorial the students are asked whether an analysis based on the articulatory and acoustic phonetic concepts previously discussed in the lectures might help determine what the politician actually said at the Iowa campaign stop. Provided with a list of relevant concepts including 'phoneme', 'stop' and 'formant', the students are tasked with identifying the steps that a forensic linguist might take to shed some light on the controversy. They are also informed that their approach and possible solution to the authorship problem needs to be accessible to a non-linguistic audience, as the application of linguistics to forensic contexts is only one step to analysing language evidence. The other step is the ability to convince the layperson of the value of linguistics in such analyses. This second step is especially challenging because, as forensic linguist Malcolm Coulthard and others have observed, the layperson—whether judge, jury member, attorney, witness or defendant—generally have strong opinions about a given piece of language evidence (Coulthard and Johnson, 2007).

Working in groups in the face-to-face tutorial—or individually on the distance online forum—students tease out, present and discuss a number of sequenced steps for the problem at hand. In the past three years, the sequential steps proposed by the students have roughly been as follows:

> *Step 1 – Identify which part of the controversial statement*
> *should be analysed*
> This is a crucial starting point because, in the case at hand, students realise that an acoustic analysis of *black* versus *blah* needs to take into account the fluidity of the sequence in which the words have been uttered, meaning that the juncture between the final sound in *black* or *blah* needs to be examined in relation to the initial sound in *people*.
>
> *Step 2 – Identify how many sounds comprise the sequence*
> This and the following two steps illustrate that the authorship problem is one of sounds, not letters. Although self-evident for linguists, this observation can be initially perplexing for the layperson schooled in the alphabet from an early age.
>
> *Step 3 – Identify each segment as C or V*
>
> *Step 4 – Identify each C or V as a specific sound*
> Students have learned about IPA prior to the tutorial but not all students will use IPA here. Instead, some students draw on descriptions like

'the *trap* vowel' or a '*k* sound'. This less than precise description is not crucial for the outcome of the activity and, interestingly, reflects the way in which a forensic linguist might present articulatory information to a lay audience, at least initially.

Step 5 – Plot the statement on a spectrogram and analyse the sequence identified above
Students use PRAAT for this step. The acoustic software has already been introduced and practised in the previous week's lecture.

Step 6 – Obtain exemplar recordings for comparison
Although this step has been identified and would most likely form part of a forensic linguist's analysis of this particular authorship problem, it is not practical because students do not have access to a speaker of American English whose sociolinguistic background is the same as that of the politician in question.

Once the steps have been discussed, the students embark on a general discussion of what they can see in the spectrogram of *black/blah people*. Significantly, the actual spectrogram of the controversial statement is inconclusive and contradictory. Namely, there is a fair amount of silence, which would be indicative of the juncture between the adjacent stops in *black people*. However, the vowel formants in the first word do not clearly suggest the low front vowel /æ/ like that in *black*. In fact, the formants are more indicative of the diphthong /aɪ/, which suggest the politician uttered [blaɪk]. Although the YouTube video that the students started with clearly leads a person to believe they are hearing *black people* as opposed to the politician's claim of *blah people*, it is quite possible that, as he claims, the politician did change his utterance, if ever so slightly.

This tutorial activity reveals that forensic linguistic data can be indecisive, in spite of the rigours of analysis. But what is most important for the students here is not the outcome of the analysis *per se* but rather the steps that the analysis comprises and the linguistic concepts in which those steps are grounded. But beyond the learning of the specific knowledge and processes involved in the task, the video anchor helps the students see their value in making sense of and taking an educated position on a high-stake, real real-life context that has attracted much attention, as highlighted by the footage itself and the reactions to it from the media and the YouTube community. Linguistic data 'anchored' in this way allows students to step away from their readings

and apply what they are learning to real-life situations whose understanding depends on specific linguistic concepts and tools.

4.2 Assessment task: Language planning

Language and Society in New Zealand is a course designed to teach the fundamentals of sociolinguistics with specific reference to the local context. It includes units on NZ English and te reo Māori, and relies largely on New Zealand-based research and examples as an initial platform to introduce the key elements of the subfield before exploring their applications to international contexts. As with other linguistics papers in the programme, this course makes use of the principles of anchored instruction during face-to-face lectures and tutorials, as well as in the materials designed for distance teaching. In addition, in this course, elements of this approach are found in the design of the students' assessment tasks.

One of these tasks, centred around the unit on language planning, requires the students to imagine they are acting as language planning consultants to advise the Angolan government on which language or combination of languages should feature in a hypothetical new Angolan national anthem. The main aim of the assignment is to assess the students' ability to recall and appropriately apply the sociolinguistic knowledge they have gained during the term to propose a rational and defensible solution to the kind of real-world problem a career linguist might be called upon to solve. Ultimately, the problem of language choice in a national anthem replicates previously introduced content relating to official languages and anthem languages in New Zealand: by encouraging the students to recollect such content and use it to solve the Angolan problem, the task is aimed at building an understanding of language planning principles and their applications beyond the local context.

The assignment is presented to the students as consisting of three parts. In Part A, the students are given the assignment instructions (in written form) and the instructor explains the main task to the class. To help the students begin their exploration of the linguistic situation of Angola (which they will further on their own and which will form the basis of their recommendation to the Angolan government), the students are shown a recent documentary about Angola. The documentary, which is available on YouTube, offers some basic information about the African country, its recent history and its languages.

The internal students are shown parts of the documentary in class during a tutorial session, while the distance students are given the video's URL and directed to view it online. As part of the introduction to the documentary,

Figure 2: *Made in Angola* (YouTube, n.d.)

all students are encouraged to take notes during the viewing, to record information that may be of use in forming a decision about the language(s) to be used in the new anthem.

In part B, the students are made to work together to discuss the content of the documentary using a number of prompts designed to encourage a discussion on what issues should be considered in making a decision regarding the anthem. In the classroom the discussion takes place in small groups, while distance students make use of a specific online discussion forum to post their thoughts and comments. In the exchanges that follow, key concepts are recalled, relevant information is extracted, preliminary opinions emerge and suggestions for further research and possible information sources are identified and shared.

Finally, in Part C, the students need to work individually to further their research and apply their knowledge of sociolinguistics to the information they have gathered by writing a formal report that presents and justifies their final recommendation.

While the nature of the anchor in this task makes it less likely for the students to already be familiar with it, its use as a way to introduce a real-life context and as a source of key data reflects the anchored nature of the assignment, and particularly of Part A and B, which are not in themselves assessed, but which are crucial to preparing the students for the task in part C.

In part A, the assignment's instructions and the documentary work together as the anchor through which the students are presented with an example of the 'kind of problem that experts may encounter and the knowledge that these experts use as tools' (CTGV, 1990:3). This part of the assignment reflects the importance that authenticity (of teaching content, contexts and tasks) is given in anchored instruction as a way to avoid rote or inert learning, whereby learners are able to recall knowledge when asked to, but unable to use it spontaneously in problem solving. The imaginative component of the task, which requires students to adopt the persona of a language planner (i.e. an expert), contributes an important element of stimulation and interest, while at the same time minimising the impact of potential emotional barriers between the student as a novice and/or outsider and the field, helping learners to access and apply knowledge of a linguistic nature within an unthreatening environment, while at the same time preserving the authenticity of the task.

In terms of the advantages of this task for the specific audience the course is designed to teach, particularly noteworthy is the possibility to lead students with little or no linguistics background to engage in a task that illustrates some of the discipline-related theoretical and practical aspects while at the same time drawing attention to linguistics and its applications to real-world contexts. Because of the argumentative nature of the assignment, which gives students the freedom to come to their own solution by constructing their own argument for their final choice of language(s) to be included in the anthem, the task tends to appeal to students who approach linguistics from the main perspective of other disciplines, and who are likely to use such perspectives to enrich their work. In other words, by highlighting the complex nature of sociolinguistics and its connections with related fields of knowledge and/or experience, the task helps showcase the significance of linguistics, particularly to non-majoring students, as well as to students who are completely new to the subject.

Part B of the assignment is designed as an opportunity for the students to collaborate with their peers in the preliminary stages of the problem-solving part of the assigned task. This phase is again well aligned with the principles of anchored instruction, as 'one of the goals of anchored instruction is to

help create environments that are conducive to cooperative learning' (CTGV, 1993:58). In particular, this phase, in which the students come together to brainstorm on how to use the information they collected from the documentary to begin shaping their recommendation and the supporting argument, is intended as a means to ease the students into the problem, which, partly because of its realistic nature, can be perceived as challenging and even threatening. The benefits of collaboration are particularly high for students who are completely new to linguistics and who might not yet have a firm idea of what linguistics is, as within the safer and more supportive environment of small group work, they are given the opportunity to take control of their own engagement with the task, feeling supported in their transition into the role of expert. Students who are new to the subject but have some academic grounding in other fields might use their knowledge as an entry point to the discussion and as an opportunity to contribute a fresh perspective on the problem. It is not uncommon for students' discussions at this stage to take interesting and unexpected turns resonating with ideas and terminology associated with fields such as marketing, psychology and human development. In line with the tenets of the social constructivist perspective underpinning anchored instruction, the role of the teacher as a facilitator during this stage is particularly important, not only to help learners sift through the available information and to emphasise possible links to relevant sociolinguistic concepts, but also to highlight the value of students' contributions by validating cross-discipline and cross-contextual connections. In the case of the distance learners' discussion, the same benefits extend to students who have extensive life and/work experiences, which, when contributed to the discussion, can encourage engagement and stimulate learning, enhancing students' motivation by creating feelings of self-worth and self-confidence as well as contributing to a sense of community that we know to be crucial to learning in distance mode.

5. Conclusion

As the examples above illustrate, within our specific context the principles of anchored instruction can offer useful affordances for teaching linguistics to non-linguists to meet the learning needs and interests of a complex and varied student population across two different modes. In our experience, consistently supported by positive feedback from our students, there are significant benefits in linking linguistic content with real-world issues and scenarios through

an anchored instruction approach. Especially when introducing linguistic principles and/concepts to students who are not pursuing linguistics as a major or a minor, the use of realistic anchors that connect with their other academic interests and personal experiences can enhance the immediacy, interest and relevance of linguistics as a subject for individuals from a wide range of backgrounds and walks of life. As with all teaching content and activities in our linguistics courses, we view authenticity as an absolute necessity in order to construct effective—albeit non-traditional—learning environments that help our students develop their knowledge of linguistics and their identity as a linguist (if that is what they aspire to) while at the same time providing effective training in communicating linguistics to laypeople, a skill which we view as particularly important for our students, given the wide variety of professions and careers they will be likely to pursue in the future.

References

BC. N.d. Retrieved from http://www.massey.ac.nz/massey/learning/programme-course-paper/programme.cfm?prog_id=93330

Bransford, John D., Robert S. Sherwood., Ted S. Hasselbring., Charles K. Kinzer and Susan M. Williams. 1990. 'Anchored instruction: Why we need it and how technology can help'. In Don Nix and Rand Spiro (eds) *Cognition, education, and multimedia: Exploring ideas in high technology*. Hillsdale, New Jersey: Lawrence Erlbaum Associates. 115–141.

Brown, John S., Allan Collins and Paul Duguid. 1989. 'Situated cognition and the culture of learning'. *Educational Researcher, 17(1)*, 32–41.

Cognition and Technology Group at Vanderbilt University (CTGV). 1990. 'Anchored instruction and its relationship to situated cognition'. *Educational Researcher,* 19(3), 2–10.

Cognition and Technology Group at Vanderbilt University (CTGV). 1993. 'Anchored instruction and situated cognition revisited'. *Educational Technology* 33(3), 52–70.

Collins, Allan, John S. Brown and Susan E. Newman, 1989. 'Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics'. In Lauren B. Resnick (ed) *Knowing, learning, and instruction: Essays in honor of Robert Glaser* Hillsdale, New Jersey: Erlbaum. 453–494.

Coulthard, Malcolm and Allison Johnson. 2007. *An Introduction to Forensic Linguistics: Language in Evidence*. London: Routledge.

Herrington, Anthony and Jan Herrington. 2007. 'What is an authentic learning environment?' In Lawrence A. Tomei (ed) *Online and distance learning: Concepts, methodologies, tools, and application*s. Hershey, Pennsylvania: Information Science Reference. 68–76.

Lave, Jean and Etienne Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation.* Cambridge: Cambridge University Press.

Made in Angola. N.d. Retrieved from
https://www.youtube.com/watch?v=1gdrQ7MV9Qs

McMenamin, Gerald R. 2002. *Forensic Linguistics: Advances in Forensic Stylistics*. Boca Raton, Florida: CRC Press.

Rick Santorum said blah or black. N.d. Retrieved from
https://www.youtube.com/watch?v=VR-ceU0oXBo

Whitehead, Alfred N. 1929. *The aims of education*. New York, New York: New American Library.

LANGUAGE TEACHERS' NEED FOR LINGUISTICS

Una Cunningham: *School of Teacher Education, University of Canterbury* *<una.cunningham@canterbury.ac.nz>*

Abstract

This paper argues that language teachers need to have knowledge of the linguistics of the language they are teaching. It explores the relationship between implicit and explicit knowledge of the target language and the consequences of this relationship for learners and teachers. Previous research has shown that a focus on form and the teaching of explicit knowledge about the target language are useful for some kinds of language learners. Modern language courses at university level form the content knowledge that future language teachers base their teaching on, yet they may not systematically cover the linguistics of the language in question. It is argued that a basic understanding of linguistics and some formal study of the structure of the target language will allow creative language teachers to better help the language learners they teach to see that the target language is organised in particular ways and make connections with their own language(s). This is a stipulated Achievement Objective for the first levels of the Learning Languages learning area in the New Zealand Curriculum. Language learners may not be served by being taught to talk about the sentence structure and phonology of their target language, but it is argued that their teachers need this knowledge to better identify and respond to the needs of the learners.

1. What does it mean to teach a language?

The target of contemporary foreign language teaching is to help learners to learn to communicate in their target language. Despite centuries of dedicated effort, language educators have been unable to come up with a really easy or even halfway effective approach to language teaching. This is all the more astonishing given the resounding success met by almost all who embark on the acquisition of their first language, sometimes with minimal support from their environment. Contemporary communicative language education in the language classroom is carried out primarily by inducing learners to perceive a need to actually communicate in the target language and by teaching them learning and communicative strategies to compensate for gaps in their knowledge.

This might seem obvious to the modern reader, but it is a total turnaround from the targets and teaching of just a decade ago, where learners who had studied a language for five-ten years were skilled in the completion of grammar tests and vocabulary tests, but unable to communicate spontaneously in the target language. Previously the target was for learners to learn about the target language and they were taught grammatical rules and drilled in their application and given lists of vocabulary to learn. Their knowledge of these rules and the vocabulary items they had been asked to learn were what was assessed. In large parts of the world, English language is still taught and assessed in this way (cf. e.g. Le, 2011). Learners from these contexts may not have been required to speak their target language before university level study, if then. Reading and listening comprehension, grammar and vocabulary were much more important to these learners and their teachers than speaking, writing and interaction.

Curriculum reform has changed this in many parts of the world. In New Zealand this began in 2007 when Learning Languages was introduced as a learning area for schools, and was refined in 2011 (New Zealand Curriculum, 2010) and earlier in most of Europe thanks to the introduction of the Common European Framework with functionally defined levels of proficiency (Council of Europe, 2001). The New Zealand Curriculum, which is heavily influenced by the Common European Framework (CEF), suggests that 'Learning a new language provides a means of communicating with people from another culture and exploring one's own personal world' (New Zealand Curriculum, 2010). This is all about using the language; there is no suggestion that a

competent speaker needs explicit language knowledge. The CEF's B1 level global scale puts it like this: a language user at level B1:

> ...can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics, which are familiar, or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans (Council of Europe, 2001).

The NZ Curriculum Level 7 corresponds to B1 on the CEF and will correspond to learners sitting for the NCEA Level 2 qualifications in e.g. French. The Achievement Objectives here are expressed separately for Communication, Language Knowledge and Cultural Knowledge, though only Communication is assessed.

Communication

Students can use language variably and effectively to express and justify their own ideas and opinions and support or challenge those of others. They are able to use and identify the linguistic and cultural forms that guide interpretation and enable them to respond critically to texts. (Adapted from *Common European Framework for Languages*, Global Scale Level B1: Independent User; Council of Europe, 2001.)

In selected linguistic and sociocultural contexts, students will:

- ☐ Communicate information, ideas, and opinions through increasingly complex and varied texts.
- ☐ Explore the views of others, developing and sharing personal perspectives.
- ☐ Engage in sustained interaction and produce extended text.

Language knowledge

Students will:

- ☐ Analyse ways in which the target language is organised in different texts and for different purposes.
- ☐ Explore how linguistic meaning is conveyed across languages.

Cultural knowledge

Students will:

☐ Analyse ways in which the target culture(s) is (are) organised for different purposes and for different audiences.
☐ Analyse how the use of the target language expresses cultural meanings.

Note that there is no suggestion that a competent speaker of the target language needs explicit language knowledge in order to communicate in the target language, just that learners should notice differences between the target language and languages familiar to them. The role of 'noticing' in language learning will be discussed below..

2. Implicit and explicit language knowledge

Learning a language does not mean the same thing to everybody, and it does not mean the same thing now as it did in the past. As mentioned above, previously, and still in some contexts, explicit knowledge about the language is what is assessed, often in grammar tests, vocabulary tests or being asked to translate sentences into and out of the target language. Implicit language knowledge, being able to actually use the language, is, however, central to what we currently mean by learning a language. One question concerns whether explicit teaching about the language leads to implicit knowledge of the language and the ability to use it to communicate.

To be able to communicate in the target language with any kind of spontaneity, learners need to develop their ability to put together functional utterances in such a way that other speakers of the language understand what they mean. This is not to say that the learners need to have explicit knowledge of the structure of the language. They do not need to understand or be able to talk about sentence structure or collocations or the phonotactic constraints of the target language, although they will be required to use these reasonably well to express themselves adequately. Compare this to the competent driver who does not need to be able to explain how the engine works, with gears and pistons, just to skilfully use the foot pedals and other controls.

Depending on the expected learning outcomes, explicit language knowledge, that is metalinguistic competence, actually being able to talk

about the language structures, and compare them with the structures of known languages, may or may not be part of the syllabus. And even if explicit knowledge is part of the language-learning syllabus it may not be assessed, as is the case for the New Zealand Curriculum Learning Languages area:

> The achievement objectives in the Communication strand provide the basis for assessment. The two supporting strands, Language knowledge and Cultural knowledge, are only assessed indirectly through their contribution to the Communication strand. (Ministry of Education, 2014)

There are many ways to deliver explicit instruction to help learners develop explicit language knowledge. Several studies (e.g. Haight et al. 2007) compare the teaching of the structure of a language form deductively (rule explanation before use of the targeted grammar feature in exercises) and inductively (practice before rule presentation). Haight et al. found that college students learning French using an inductive approach had better grammatical knowledge of structures. This, of course, says nothing about their ability to actually use the language. The question they addressed was: 'What is the most effective approach to teaching grammar in a foreign language classroom?' (Haight et al. 2007: 289).

De Graaf & Housen (2011: 737) contrast implicit and explicit forms of form-focused instruction, and relate these to Long's (1991) distinction between *focus-on-form* instruction that 'overtly draws student's attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication' (Long 1991:45-6) rather than the *focus-on-forms* which are decontextualized, such as random sentences used to exemplify structures, or lists of vocabulary words in isolation to be memorised.

Teaching grammar is not, of course, the same as teaching language, and if it is actually the learners' grammar knowledge that is assessed, then both implicit and explicit teaching may well be useful. Given that implicit knowledge (that is the tacit knowledge that a language user has of the language being used) is the target, the important question is whether explicit teaching leads to implicit knowledge. N. Ellis (2005) attempts to answer the question of how explicit knowledge about language affects implicit language knowledge of language. He refers to, among others, Krashen (1985) who separated acquisition and learning and claimed that learned language could not be used 'on the fly', and Paradis (1994) who claimed that explicit knowledge does not become implicit knowledge, but states that explicit and implicit language knowledge do interact and that empirical research in the preceding 30 years showed that

'language acquisition can be speeded by explicit instruction' (Ellis 2005:307). In fact, results are quite mixed, but do suggest an effect from explicit form-focused instruction.

Other work strives to apply in language education insights gained from decades of research in Second Language Acquisition (SLA). Traditionally there has been limited communication between linguists involved in SLA and language educators. R. Ellis (2012) gives an excellent review of empirical classroom-based studies looking at the application of these ideas to real language teaching, demonstrating the very mixed outcomes of classroom-based studies in general. A recent issue of *Applied Linguistics* (R. Ellis, 2015) is entirely dedicated to narrative reviews and meta-analyses of the efficacy of language instruction. This includes papers on the acquisition of second language grammar (Dekeyeser & Prieto Botana, 2015; Shintani, 2015) and on the effectiveness of second language pronunciation instruction (Lee et al, 2015; Thomson & Derwing, 2015). The debate is clearly ongoing.

In a meta-analysis, Norris & Ortega (2000) found explicit instruction (where metalinguistic rules were explained to learners) to be more successful than implicit instruction (where learners were exposed to forms but their attention was not drawn to the forms and they were not explained). Ortega (2009) also asks if explicit teaching (grammar instruction) helps, and explores the relationship between implicit (subconscious) and explicit (conscious) learning in connection with implicit (inductive or indirect) and explicit (deductive or overt) instruction. She acknowledges that the body of evidence suggests that 'instructed learners progress at a faster rate, they are likely to develop more elaborate language repertoires and they typically become more accurate than uninstructed learners' (Ortega, 2009:139). She illustrates this claim with reference to studies of learners' mastery of the syntax and morphology of their target languages. She goes on to point out the problems inherent in SLA constructs such as *interlanguage* and *target language* since they assume a monolingual native speaker model (and target, cf. Cunningham, 2009).

There are a number of reasons why instruction might not work well. One is that discussed above, the uncertainty of learners being able to 'convert' explicit knowledge to implicit knowledge. Another is whether or not some parts of the target language are teachable at all to the learners at hand and whether they are learnable by them. This might be because the learners are not 'ready' for particular features in terms of having mastered preceding stages of learning (Peinemann, 1989) or because of specific phonotactic features or

phonological rules of the first language that are getting in the way of the target language features (Cunningham, 2013).

Interface/non-interface hypotheses posit that explicit knowledge can or cannot become implicit knowledge. De Graaff and Housen (2011) outline the three theoretical possibilities offered by the no-interface, weak-interface, and strong-interface hypotheses and their consequences for language teaching. They then relate these to current thinking on meaning-focused instruction, such as task-based language teaching, a strong form of meaning-focused instruction. Their conclusion is that research on accuracy suggests that instruction helps, but that studies of learner fluency are inconclusive, mostly because of the difficulty of studying extensive language use rather than controlled use.

3. Contemporary language education

The problem, and at the same time the solution to the problem, is that learners seem to need to actually use the language in order to learn it. The conundrum inherent in this is that learners need to have some target language vocabulary and structures to start producing utterances so that they can develop proficiency. Like infants acquiring their first language or languages, however, learners can begin honing their receptive skills before they venture to production. A difference is that while the infant spends months listening before needing to communicate through speech (though they give it their best shot), the older learner can easily move from getting input, that is hearing or reading target language items, to reproducing them with communicative intent within seconds of first exposure. The question examined here is how explicit learning applies to the learning of an additional language and how teachers can best help learners to achieve their goals. N.Ellis put this quite clearly:

> Some things we just come able to do, like walking, recognizing happiness in others, knowing that *th* is more common than *tg* in written English, or making simple utterances in our native language. We have little insight into the nature of the processing involved – we learn to do them implicitly like swallows learn to fly. Other of our abilities depend on knowing how to do them, like multiplication, playing chess, speaking pig Latin, or using a computer programming language. We learn these abilities explicitly like aircraft designers learn aerodynamics. (N. Ellis, 1994: 1)

As Long (2009) made clear, there are important differences between the conditions of first and second language acquisition and learning as well as important similarities. Older learners need to have exposure to input in the target language. Finding input for the learners is fairly easy in contexts where the target language is a main language, as in the case of migrants learning English in New Zealand, but more challenging in contexts where the target language is not the main language used, as is the case in the teaching of foreign languages in schools. The problem of appropriate input has been part of language teaching at least since the work of Comenius (1657). More recently these ideas were restated by e.g. Krashen who insisted on the importance of comprehensible input in his Input Hypothesis first published in 1977. He even went as far as to suggest that comprensible input is all you need in order to be able to learn a language, and that grammatical explanations were, at best, a waste of time.

While other aspects of Krashen's work have been criticised and later refined, debated and rejected by some (e.g. Cook, 2013), the notion that learners need input is not questioned. It is, however, generally held that input, even comprehensible or meaning-focused input is not enough, and that learners need opportunities for output, that is a chance to actually use the language in a meaning-focused way. Swain (1985) formulated this as the Comprehensible Output Hypothesis, and developments of this (e.g. Swain, 1995; Swain & Lapkin, 1995) have been extensively studied and put to the test in empirical classroom-based studies. Gass (2003) further examines the role of input and output, including modified input (such as might be used to address less proficient interlocutors) and the kind of positive and negative evidence about the nature of the target language the learner can glean from the available input, some of which is in response to the learner's own language production, such as requests for clarification or corrective feedback.

Gass (2003) goes on to consider the importance of interaction for language learning and traces the recognition of this from Wagner-Gough and Hatch (1975) via Long (1980) to full acknowledgement of the role of conversation in language learning as expressed in Long's Interaction Hypothesis (1996). Long is also celebrated for his work on focus on form (Long, 1991) which distinguishes between teaching approaches where form is an organising principle of a language syllabus and meaning-focused or communicative approaches where focus on form happens on a just-in-time basis. This has been developed further, e.g. by R. Ellis et al (2002) and Loewen (2015).

Laviosa (2014) discusses the role of translation in the history of language

teaching. It has fallen in and out of favour as the priorities of language teaching have moved from using the grammar-translation method for gaining a reading knowledge of a language like Latin to developing oral proficiency in a living modern language where the learner can reasonably aspire to communication with monolingual and other speakers. Translation can be used in language learning with or without explanation of structural differences between the learner's first language and the target language. Laviosa, (2014: 28) points out the value of translation for developing metalinguistic competence, and translation is still fairly common in language classrooms where grammar translation still is set against teaching influenced by the Direct Method exemplified by Berlitz who banned translation in the language classroom (Laviosa, 2014: 141).

Gass (2003) also considers the importance of attention or noticing, tracing the debate on whether learning without awareness is possible through the nineties and beyond. Schmidt's work (e.g. Schmidt, 1990) on consciousness in language learning has been very influential, although strong claims were nuanced in later work, suggesting that noticing might not be entirely essential for learning to happen. Interest in the role of attention and consciousness continues. Truscott (2014) problematized the notion of noticing and the distinction between conscious and unconscious processes in language learning and pointed out that different views of the importance of consciousness have lain behind ideas about language teaching throughout its history.

Even since the introduction of Communicative Language Teaching (CLT) with its focus on meaning-focused activities onto the stage, actual classroom practices have varied. Sometimes there is a zero-grammar approach, as inspired by Krashen (1985) and his thinking about the Natural Approach which aspired to emulate first language acquisition (e.g. as described in López Rama & Luque Agulló (2012: 181) where 'there is still a belief that the teaching of grammar might be harmful for communicative competence'. Communicative Language Teaching has become mainstream. The weak interpretation, of Communicative Language Teaching, Present-Practice-Produce (PPP) is sprung, as López Rama & Luque Agulló point out, from Structuralist-Generativist thinking about the nature of language. Cook (2008: 265) refers to it as 'the mainstream EFL style'. It is still organised by structures (the Present part) and grammar drills (Practice) and then, at least in theory, sets the learners free of the bonds of controlled language into a Produce phase. Unfortunately, it seems that the lesson is often over by then, and so the learners are never asked to actually use the language for non-controlled language

production (e.g. Le 2011). The strong version of CLT merges with Task-based Language Teaching (TBLT), which includes a focus on form, often presented inductively, as and when the learners need it for their communicative purpose.

Nation's work (e.g. Nation, 1996, 2007) identifies four *strands* of language education: meaning-focused input, meaning-focused output, language-focused learning and fluency development. These cannot be directly mapped to R. Ellis' ten (later eleven) principles of instructed second language instruction, which he identified in his meta-analysis of studies of successful language teaching. R. Ellis's principles are reproduced verbatim in material from the New Zealand Ministry of Education for language teachers (R. Ellis, 2005). In 2002, R. Ellis held that Second Language Acquisition research has little concrete advice to offer to language teachers or to language teacher students:

> [T]here is no agreement as to whether instruction should be based on a traditional focus-on-forms approach, involving the systematic teaching of grammatical features in accordance with a structural syllabus, or a focus-on-form approach, involving attention to linguistic features in the context of communicative activities derived from a task-based syllabus or some kind of combination of the two. Nor is there agreement about the efficacy of teaching explicit knowledge or about what type of corrective feedback to provide or even when explicit grammar teaching should commence. (R. Ellis, 2002:209).

Nonetheless, just a few years later, R. Ellis (2005) introduced his ten principles for instructed language learning and these also appeared in an extended form in a report produced for the NZ Ministry of Education in 2005 (R. Ellis et al, 2005), which led to a summarized and accessible (to busy pre-service and in-service teachers) version of this meta-analysis (R. Ellis, 2008). These principles and Nation's Four strands have been seized eagerly by language teachers and teacher educators, as, at least ostensibly, giving some concrete research-informed advice to the classroom teacher that can be used to inform lesson planning. Both R. Ellis and Nation advocate a focus on meaning tempered with some focus on form.

4. Learner preferences

Sewell (2004) points out that translation activities can be more attractive to students as they give them assurance that they have fully understood how to use the target language structures. In a communicative approach to language

learning, on the other hand, the teacher may not correct all errors, and students may be unsure of whether or not they are using the correct forms. Sewell (2004: 153) lists the needs felt by many language learners for confidence and self-esteem, not to lose face, to be rewarded, for certainty, closure and autonomy as well as needs arising from introversion. She mentions the advantages that a 'closed-ended' activity like translation, with a verifiable outcome, the translated text, has over an open-ended communicative language task like carrying out a transaction in a role play of a chemist's shop. 'There is nothing concrete to measure your performance against, since you are not actually in a chemist's shop' (Sewell, 2004: 158).

The same can be said of the security offered by explicit grammatical explanations. Learners who are exposed to an inductive approach to focus on form where they are invited to find patterns in target language material, will be frustrated if they are left without a wrap-up deductive explanation, a key to the exercise they have been working through. This may not actually be important for their learning. Jean and Simard (2014) found that while the junior high students they studied expressed a preference for a deductive approach, their preference did not appear to be related to the success of their learning.

A good deal is now known about individual differences between language learners. While attempts to teach the strategies and practices of 'good language learners' to other learners have been disappointing (Rubin 1975), there are other insights that can be gained from such studies. De Graaff and Housen (2011) summarise this area with reference to studies of the effects of learning style, age, language-learning aptitude, proficiency level and other variables on the effectiveness of instruction. It seems likely that while explicit teaching helps for some learners, for others it is just confusing.

5. Teacher knowledge about the target language and learners' first language(s)

While language schools and education systems in many parts of the world try to recruit native speakers of English to teach the language to learners, most language teachers have learned the target language at some point. Teachers who are themselves first-language speakers of the target language are in a different position than these teachers. Regardless of whether or not the first-language speaker is a qualified language teacher, and regardless of whether or not they speak the language(s) of their learners, they will not have the same

experience as a teacher who has learned the target language as an additional language. They will have their native-speaker intuitions (implicit language knowledge), but be unequipped to justify or communicate them. If they have never themselves learned an additional language, they are unlikely to have ever considered the situation their students are in.

Phillipson (1992) is sceptical of the value of native speakers as teachers, even as models. He speaks of a 'native speaker fallacy' (Phillipson 1992:195) and argues that fluent, idiomatically appropriate language is not impossible for non-native teachers to learn, but that the insight teachers have into the processes of language learning and the structure and usage of a language must be learned. He goes on to say that the capacity to analyse and explain language definitely has to be learned, even though he suggests that it need not be taught. He writes: 'the untrained or unqualified native speaker is potentially a menace' (Phillipson, 1992: 195). He refers to a UNESCO report from 1953 warning that being a native speaker is not a qualification for teaching and suggests that the ideal language teacher is a successful language learner who has 'a detailed acquaintance with the language and culture of the learners they are responsible for' (Phillipson, 1992: 195).

Another complication that Phillipson points out in the case of English is the fact that English is far from homogenous. A local indigenized variety of English may exist alongside a standard language norm, and a native-speaking teacher will necessarily bring yet another English to the learners. Phillipson mentions the deep conflict at that time between proponents of an exonormative model, such as on the one hand Quirk (1990) who believed that the 'leading English-speaking countries' had a right of interpretation of how English should be taught and on the other hand, Kachru (1991) with his endonormative thinking, valuing established periphery varieties of English.

A teacher with little or no linguistic training will find it difficult to understand the challenges experienced by the learners. In the case of pronunciation, if the teacher has explicit knowledge of the phonotactics and phonological rules of the learners' first language(s) and how they differ from those of the target language, they will better understand why particular systematic features of the learners' pronunciation appear. The teacher is then in a position to point out to the learners the differences in form between the target language and the first language(s). If the teacher does not share the learners' language background, they will also benefit from explicit linguistic knowledge of their own language(s), so they will understand differences between their own speech and that of the learners. This is the case even if

the target language is the teacher's first language; linguistic knowledge of the teacher's own variety as well as of standard varieties of the languages concerned is very useful for the teacher's understanding of what is difficult for the learners and why. This is not, of course, to say that this needs to be taught to the learners.

A linguistically naïve teacher who has learned the target language will be better equipped in some ways to teach the language than a first language speaker of the target language, but still unable to talk about the language structure in a useful way. In a homogenous class where the learners share a first language, knowledge of the phonology, morphology and syntax of both the target language and the learners' first language makes it possible to use the learners' implicit knowledge of their language to notice similarities and differences between the languages.

I am certainly not advocating more focus on forms or the use of form as an organising principle for language teaching. I believe firmly that focus on meaning should be the rule with formal matters being discussed as they arise for the communicative purpose of the learners. However, for any focus on form to take place, even incidental focus on form at the point of need, the teacher will need to have some level of awareness of the syntactic and phonological organisation of the target language. If the teacher is a linguistically untrained native speaker, s/he is unprepared to help the learner become aware of structure. If the teacher is a fellow learner of the target language, with years of study of the target language and possibly its literature(s) and culture(s) behind them, but no actual explicit linguistic knowledge of the target language, as is the case for many Modern Language graduates in New Zealand and elsewhere, s/he will not be able to help the learners acquire explicit language knowledge. If the teacher does not have knowledge about the learners' language(s) as well as of the target language, there can be no contrastive analysis. While the strong form of Contrastive Analysis has been shelved as an approach to language teaching because of its failure to predict points of difficulty for learners other than in superficial ways (Ortega 2009), it is not without value. Knowing a little about articulatory phonetics and voice onset time, for example knowing that Spanish word-initial [t] in *te* is a voiceless unaspirated apico-dental stop and that this is other than the voiceless aspirated lamino-alveolar [t] in the English word *tea*, will be helpful for the person teaching English to Spanish learners as well as for the person teaching Spanish to English learners, regardless of the language(s) spoken by the teacher. The learners do not need this metalinguistic knowledge – for them it is enough to approximate a native place of articulation

for the Spanish teacher to say "put your tongue tip against your top teeth". The difference in VOT, the Spanish initial voiceless stops being unaspirated, seems to be much harder to teach to English speakers, and will often be acquired through exposure or can be taught using noticing strategies at a later stage. But if the teachers do not have this explicit knowledge they will not be able to give accurate and helpful instruction. In fact they may give explanations based on their own half-formed notions about the difference between the students' L1 and the target language.

6. Language teachers' need for linguistics

In conclusion, there is compelling evidence in the literature that many, but not all, language learners will be helped by various kinds of form-focused instruction, be it explicit explanation of metalinguistic rules or analytical teaching to help them develop explicit language knowledge, or implicit teaching by enhanced or unenhanced exposure to problematic target language forms. This is not to say that the learners should themselves learn the linguistics of the target language, rather that the teachers should have that knowledge. A linguistically untrained language teacher lacks the wherewithal to offer this kind of focus-on-form teaching. In addition, teachers who have not been linguistically schooled are susceptible to diverse fads and methods that they happen upon and may latch onto wholly or partially. In particular, there are problems in language teacher education associated with the move from traditional grammar-laden form-focused approaches to language teaching to communicative, meaning-focused approaches. While the latter may claim to assess only communicative competence unlike the former which taught and assessed only or mostly explicit knowledge of the language, this does not mean that learners do not need form-focussed instruction from time to time. Ellis, Nation, and others whose work informs contemporary language syllabi in many parts of the world have never suggested that. An overwhelming body of empirical research insists that form-focused instruction leads to faster, more direct routes to higher proficiency (de Graaff & Housen, 2011).

Even teachers who have themselves learned the target language with a traditional form-focused grammar-laden approach are not well equipped to offer this kind of teaching unless there has also been some linguistic element in their education. By this I mean that it is not enough for an English teacher (by which I mean an EFL/ESL/ESOL teacher) to have

studied English pronunciation as a language learner – some study of English phonetics and phonology is also necessary. Similarly, being able to excel in an English grammar test is no substitute for at least some knowledge of English morphology and syntax at a more abstract level. The *learners* do not need this knowledge, but to teach e.g. pronunciation efficiently, the *teacher* needs explicit knowledge of phonetics and phonology. To teach grammatical accuracy, the teacher needs some knowledge of morphology and syntax. The learner, however, can happily be spared learning this conceptually difficult material. It will not enhance their ability to *use* the target language, just to *talk about* it, which is not part of the curriculum. Language teachers need to have studied the structure of at least one language – if they have done that they will be empowered to learn about other languages their learners are dealing with, including the learners' first language. Only then are teachers prepared to carry out research-informed language teaching.

References

Comenius, Johan. A. 1657. *Opera didactica omnia*. Pragae: Academia Scientiarum Bohemoslovenica.

Cook, Vivian. 2008. *Second Language Learning and Language Teaching*. Retrieved from http://www.eblib.com.

Council of Europe. 2001. *A Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Cunningham, Una. 2009. 'Models and targets for the pronunciation of English in Vietnam and Sweden.' *Research in Language*, 7, 113–128.

Cunningham, Una. 2013. 'Teachability and learnability of English pronunciation features for Vietnamese-speaking learners.' In Ewa Waniek-Klimczak and Linda R. Shockey (eds) *Teaching and Researching English Accents in Native and Non-native Speakers*: Berlin Heidelberg: Springer Verlag. http://dx.doi.org/10.1007/978-3-642-24019-5. 3–14.

De Graaff, Rick and Housen, Alex. 2011. 'Investigating the effects and effectiveness of L2 instruction.' In Michael H. Long and Catherine J. Doughty (eds) *The Handbook of Language Teaching*. Oxford: Blackwell Publishing. 726–755.

DeKeyser, Robert. 2003. 'Implicit and explicit learning'. In Catherine J. Doughty and Michael H. Long (eds) *The Handbook of Second Language Acquisition*, Oxford: Blackwell Publishing. doi: 10.1002/9780470756492.ch11

Dekeyser, Robert and Prieto Botana, Goretti. 2015. 'The effectiveness of processing instruction in L2 grammar acquisition: a narrative review'. *Applied Linguistics* 36.3: 290–305. doi:10.1093/applin/amu071.

De Wit, Astrid, Patard, Adeline, and Brisard, Frank. 2013. 'A contrastive analysis of the present progressive in French and English'. *Studies in Language*, 37.4: 846–879. doi:10.1075/sl.37.4.05wit.

Ellis, Nick C. 1994. 'Introduction: Implicit and explicit language learning. An overview'. In Nick C. Ellis (ed) *Implicit and Explicit Learning of Languages*. London: Academic Press. 1–31.

Ellis, Nick C. 2005. 'At the interface: dynamic interactions of explicit and implicit language knowledge.' *Studies in Second Language Acquisition* 27.2: 305–352.

Ellis, Rod. (2005). 'Principles of instructed language learning.' *System*, 33.2: 209–224.

Ellis, Rod. 2009, *Implicit and explicit knowledge in second language learning, testing and teaching*. Buffalo and Bristol: Multilingual Matters.

Ellis, Rod. 2005. *Instructed Second Language Acquisition: A Literature Review*. Wellington and Auckland: Research Division, Ministry of Education.

Ellis, Rod. 2012. *Language Teaching Research and Language Pedagogy*. Oxford: Wiley-Blackwell.

Ellis, Rod. 2015. 'Introduction: complementarity in research syntheses.' *Applied Linguistics 36* (3): 285–289. *doi:10.1093/applin/amv015*

Ellis, Rod, Basturkmen, Helen. & Loewen, Shawn. 2002, 'Doing focus-on-form.' *System* 30: 4: 419–432.

Gass, Susan. M. 2003. 'Input and Interaction', in Catherine J. Doughty and Michael H. Long (eds) *The Handbook of Second Language Acquisition*, Oxford: Blackwell Publishing. doi: 10.1002/9780470756492.ch9

Haight, Carrie E., Herron, Carol and Cole, Steven P. 2007. 'The effects of deductive and guided inductive instructional approaches on the learning of grammar in the elementary foreign language college classroom.' *Foreign Language Annals* 40.2: 288–310.

Jean, Gladys. & Simard, Daphnée. 2013. 'Deductive versus inductive grammar instruction: investigating possible relationships between gains, preferences and learning styles.' *System* 41.4: 1023–1042.

Krashen, Stephen D. 1977. 'Some issues relating to the monitor model'. In H. Douglas Brown, Carlos Alfredo Yorio, and Ruth H. Crymes (eds.) *Teaching and Learning English as a Second Language: Trends in Research and Practice: On TESOL '77: selected papers from the eleventh annual convention of teachers of English to speakers of other languages, Miami, Florida, April 26 – May 1, 1977*. Washington, DC: Teachers of English to Speakers of Other Languages. 144–158.

Krashen, Stephen D. 1985. *The Input Hypothesis: Issues and Implications*. New York: Longman.

Laviosa, Sara. 2014, *Translation and Language Education: Pedagogic Approaches Explored*, Routledge, New York.

Le Van Canh. 2011. 'Form-Focused Instruction: A Case Study of Vietnamese Teachers' Beliefs And Practices'. Unpublished PhD dissertation. Hamilton, NZ: University of Waikato.

Lee, Junkyu, Jang, Juhyun and Plonsky, Luke. 2015. 'The effectiveness of second language pronunciation instruction: a meta-analysis.' *Applied Linguistics* 36.3: 345–366. doi: 10.1093/applin/amu040.

Long, Michael H. 1991. 'Focus on form: A design feature in language teaching methodology.' In Kees De Bot, Ralph B. Ginsberg and Claire Kramsch (eds) *Foreign Language Research In Cross-Cultural Perspective*. Amsterdam: John Benjamins. 39–52.

Long, Michael H. 1996. 'The role of the linguistic environment in second language acquisition.' In William Ritchie and Tej Bhatia (eds) *Handbook of Second Language Acquisition*. San Diego: Academic Press. 413–468.

Long, Michael H. 2009. 'Methodological principles for language teaching.' In Michael H. Long and Catherine J. Doughty (eds) *The Handbook of Language Teaching*. Oxford: Blackwell Publishing. doi: 10.1002/9781444315783.ch21

López Rama, José, and Luque Agulló, Gloria. 2012. 'The role of grammar teaching: from communicative approaches to the Common European Framework of Reference for Languages.' *Revista de Lingüística y Lenguas Aplicadas* 7: 179–192.

Loewen, Shawn. 2015. *Introduction to instructed second language acquisition*. New York: Routledge.

Ministry of Education. 2014. 'Learning languages achievement objectives'. Retrieved from: http://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum/Learning-areas/Learning-languages/Achievement-objectives

Nation, Paul. 1996. 'The four strands of a language course.' *TESOL in Context* 6.1: 7–12.

Nation, Paul. 2007. 'The four strands'. *Innovation in Language Learning and Teaching* 1.1: 1–12.

Norris, John M. & Ortega, Lourdes. 2000. 'Effectiveness of L2 instruction: a research synthesis and quantitative meta-analysis.' *Language Learning* 50.3: 417–528.

Ortega, Lourdes. 2009. *Understanding Second Language Acquisition*. London: Hodder Education.

Phillipson, Robert. 1992. *Linguistic imperialism* Oxford: Oxford University Press.

Richards, Jack C. 2008. 'Second language teacher education today.' *RELC Journal* 39: 158–177. doi:10.1177/0033688208092182.

Rubin, Joan. 1975. 'What the "good language learner" can teach us.' *TESOL Quarterly* 9.1: 41–51.

Schmidt, Richard W. 1990. 'The role of consciousness in second language learning.' *Applied Linguistics* 11: 129–58.

Sewell, Penelope. 2004. 'Students buzz round the translation class like bees round the honey pot – why?' *Translation in Undergraduate Degree Programmes* 59: 151.

Shintani, Natsuko. 2015. 'The effectiveness of processing instruction and production-based instruction on L2 grammar acquisition: a meta-analysis.' *Applied Linguistics* 36.3: 306–325. doi:10.1093/applin/amu067

Swain, Merrill. 1985. 'Communicative competence: some roles of comprehensible input and comprehensible output in its development.' In: Susan Gass & Carolyn G. Madden (eds) *Input in Second Language Acquisition*. Rowley, Massachusetts: Newbury House Publishers. 235–256.

Swain, Merrill, 1995. 'Three functions of output in second language learning.' In: Guy Cook & Barbara Seidlhofer (eds). *Principle and practice in applied linguistics: Studies in honor of H.G. Widdowson.* Oxford: Oxford University Press. 125–144.

Swain, Merrill. & Lapkin, Sharon. 1995. 'Problems in output and the cognitive processes they generate: a step towards second language learning.' *Applied Linguistics* 16.3: 371–391.

Thomson, Ron I. & Derwing, Tracey M. 2015. 'The effectiveness of L2 pronunciation instruction: a narrative review. *Applied Linguistics* 36.3: 326–344.

Truscott, John. 2015. *Consciousness and second language learning*. Bristol; Buffalo, New York: Multilingual Matters.

CAN AN ENTIRE PAPER BE PROJECT-BASED GROUP WORK?

Hunter Hatfield: *Department of English and Linguistics, University of Otago <hunter.hatfield@otago.ac.nz>*

Abstract

Most students who are linguistics majors will not become professional linguists. Therefore, teaching of linguistics should be focused explicitly upon building the general skills of a liberal education that students will most benefit from, rather than only discipline-specific material. With this in mind, the university's course on first language acquisition was reorganized so that all instruction was based around group-oriented projects. The hope was that such a course would (1) develop students' skills with group work and research planning, (2) motivate learning through frequent interaction and (3) increase student interest due to ownership of content. For purposes of comparison, the course covered similar content as a previous version that used lectures and an extended class project. Evaluation of the course is performed through comparison of assessment results across years, student feedback on an extended survey and the instructor's experiences. Within the limitations of the study, the redesigned course did show evidence of interactive learning and team-work skills while understanding of linguistic content persisted. Moreover, Honours students in the subsequent year continued to practise techniques developed in this paper.

1. Why a Course Based upon Projects?

LING 318, Child Language, is a third-year course on first language acquisition. Most students taking the class are linguistics majors, though a small number of students come from education or psychology. Despite the fact that most students are majors, the single course they must have taken before LING 318 is an introduction to linguistics, so only some basics of phonetics, phonology, morphology, and syntax can be assumed. The current paper discusses the redesign of a lecture-based 2012 version of the class into an entirely project-based 2014 version to identify pluses and minuses of such a design.

The 2012 course used a combination of lectures and a significant project. Students read Eve Clark's (2009) *First Language Acquisition* throughout the semester. This can be a difficult text for students, and so lectures were designed around providing background material and explicating difficult concepts. For the first few weeks of the course, students attended 2-hour lectures with periodic tutorials also lead by the lecturer, myself. After this introductory period, I would lecture for the first hour, and a group of students would do a presentation on the project for the second hour. The course project was to design a test that could triage 3-year-olds into two groups, one at high risk for developmental dyslexia and a second at low risk. Dyslexia was chosen as a topic for two primary reasons. First, it is a research interest of mine and so was a chance to integrate teaching and research. Secondly, many linguistics majors do not have a chance to apply their theoretical knowledge to a question that is of immediate consequence to others. By starting the project with accounts from individuals with dyslexia about its impact on their lives, the hope was to increase interest in the material.

Assessment of the instructor-led content was based upon two take-home essay tests. The goal was to measure the student's ability to integrate ideas into a coherent theory about some aspect of child language development. As such, a successful response required using concepts from multiple chapters of the text, as well as ideas from lecture and tutorial. The focus therefore was not on testing specific material but on ability to use such material to provide a well-reasoned explanation of linguistic behaviour. Students had approximately one week to write and submit their essays. The project on dyslexia initiated eight weeks before it was to be finished. The second hour of each lecture session was dedicated to presentations from student groups who had researched some part of the topic. Students presented on topics such as the experience of dyslexia, proposed causes of dyslexia, and common methods of linguistic

research with children. The project concluded with an individual research proposal that contained a literature review and methodology for how they would perform the triage test.

Even though instructor-led lectures and tutorials composed the majority of time, the project appeared to take focus for most. One student completed an Honours dissertation based upon her dyslexia project the following year. In discussion with students about how to improve the 2012 paper, one student asked, 'perhaps the whole paper could be the project?' and another student suggested making the project more real-world by presenting the results to a support group for dyslexia. Whenever the course came up in conversation, it was invariably the dyslexia project that was mentioned. None of the content about phonological, morphosyntactic, semantic and pragmatic acquisition was discussed.

There were broader motivators as well. For one, most of the linguistics majors are not future professional linguists. It is highly unlikely that they will need to know in five to ten years whether semantic features help explain word learning or whether three-year-olds can comprehend passives. Rather, they are acquiring something more general, some set of cognitive, social and practical abilities. I am not arguing that child language is not worth studying, but that child language should be taught so as to focus on those general abilities. (Barnett & Coate 2005; Conrad & Dunek 2012; Whitehead, 1929: 26; *inter alia*). The current course redesign focuses upon developing these skills using linguistics, rather than treating linguistic knowledge as the sole end with "liberal" skills as fortunate additions.

The final motivation for a course redesign came from learning theory and learning studies. The more that students are engaged (Barnett & Coate 2005; Tagg 2003) in active tasks, rather than passive (Applebee, 1996), answering a question that they asked (Entwistle 2009), the more successful the learning outcome. In order to create such a student-directed, active and engaged environment, a collaborative-learning strategy was used (Dörnyei 1997). In one review of 168 studies on learning patterns, cooperative learning was discovered to be substantially more effective than competitive or individualistic learning. Examining effect sizes from this meta-analysis, a student who would have a score at the 53rd percentile for individual learning could be predicted to score at the 70th percentile learning cooperatively (Johnson, Johnson & Smith 1998; see also Jones & Jones 2008). In the course redesign then, groups would be created whose task was to answer some question about first language acquisition. Lectures did still occur. Indeed, the

result was that just less than 50% of the class time was composed of lectures. However, the lectures were never an end-in-themselves. They were always written to serve the group work; they were primers for the primary class activity of collaborative, active group learning.

Therefore, the skills that I hoped students would take away from the paper were those of collaborative work and independent learning – learning how to learn and teach others. The topic of first language acquisition is, in a sense, incidental. It was a course with a small enough cohort to make the design practical, challenging material, and natural ways to apply the knowledge, such as with dyslexia or childhood education. The redesigned course was created so as to be comparable with the earlier, more traditional, 2012 version. However, there are limitations to this comparability. In an ideal study, for instance, independent graders should assess essays from both years. Unfortunately, the 2012 version was not taught as part of a research project. Essays were collected, marked, and returned. As such, the current essay cannot firmly establish that a group project method, for instance, is superior. Instead, it will assess what the experiences of such a method are for the students and provide some indications regarding students' linguistic and liberal learning.

2. Design of the New Course

To assess the results of redesigning LING 318, Child Language, the outputs needed to be similar to the previous incarnation. While there is certainly a place for working on one relatively narrow topic, such as dyslexia, for an extended period, I thought that students would benefit from a broader exposure to the topic than just dyslexia. In exploring existing designs of linguistics courses and research, it was common to see projects along with lectures, where a project was an application of ideas from the lectures. However, the idea for this paper was to make 'everything' project-based, using projects to learn the main theories. It was also possible to find classes with a single, semester-long project, but I was aiming for formative, staged projects where people learn how to do projects iteratively. Finally, whatever happened must include a set of activities that a larger group (likely over 20) could perform together within constraints of lecture spaces.

Ultimately, I created a course with four smaller projects and a final exam. The first project was highly scaffolded. I supplied the research questions for every group, gave instructions of how groups should function, provided the

primary text for their learning, and set out guidelines for what they should supply to each team member in the group. By the fourth project, however, the course outline for three weeks stated simply 'what you choose.' Therefore, in a sequence I removed the scaffolding of how groups should work and the questions they should ask (Spronken-Smith & Harland 2009).

The core text was still Eve Clark's (2009) textbook; the final project was still to design a method of triaging 3-year-olds at risk for dyslexia; and the final exam contained a selection of the same questions that had been used for the 2012 take-home tests. Therefore, many of the key assessment outputs were the same. However, the teaching style was quite different. Would a project-based style of learning improve those outputs while making gains in student engagement and self-directed learning?

2.1 Projects Overview

The first three projects were targeted at learning generally about child language, covering (1) speech perception and production, (2) word learning and (3) syntax, respectively. This represents a sacrifice in breadth. However, if they were going to forget specific content, then little was lost, long-term. In these three projects, students were supposed to learn the skills for the paper and much of the content. In each of the projects 1-3, I presented a large **topic question** that guided the entire project. These questions were preparatory for the questions they would eventually see on the final exam. They also gained research experience through the three projects to get ready for project 4. In the first project, they focused upon the text and reading research articles of their choosing to answer a question. In the second project, they read the text and research articles of their choosing, and provided a new research question. In the third project, they read the text and research articles of their own choosing, and provided both a research question and a methodology. Students then demonstrated (1) the skills they had acquired in Project 4, where they researched dyslexia the way they saw fit, and (2) the knowledge they had acquired about first language acquisition in the final exam essays.

2.2 Project 1: Speech Perception and Production

The focus of Project 1 was to examine how to work as a **home group** and read primary research articles. I wrote a very broad research topic. As it was so broad, I also broke it down into seven so-called **sub-questions** that, if each were answered, would provide the knowledge to answer the overall topic question. From the first week of the semester, they were put into groups (randomly).

Each group was responsible for its own answer to the topic question, using the sub-questions as keys. Each person in the group was to become an expert on their own question. For instance, one person in the group would be responsible for learning whether babbling has any relation to early words. Another person would be responsible for learning what phonetic sounds infants can perceive. They would read the text and articles to find out the answer and then come back to teach the group.

The student was not alone, however, in becoming an expert. There were four groups of 6-7 members, implying that four students were researching the same topic. These four worked to become an expert on their sub-question (such as babbling to early words), a so-called **expert group**. The expert group operated in a pairs and square arrangement (Hughes & Townley 1994). In this arrangement, they brought possible articles to a class session. A pair from the expert group would choose one of the articles and read it. They would come back and discuss the article with each other, coming to some agreement on its findings and theory. That pair would then teach the other pair in their expert group about the article. Each person in the group was now an 'expert,' and so they would go back to their larger home group and teach those members the answer to their question. All of this finished by the home group giving a 10-minute presentation answering the topic question. Critically, the expert group functioned as a partial replacement for the material that had previously been provided in lecture. Each student had three other students to help her with difficult parts of the text, as well as using outside literature to supplement that text. If a student could increase their ability to learn material and support others, that will be a long-term beneficial skill. They practise becoming experts.

During all of this, I had two main tasks. The first was to prepare students to be experts on their sub-question. I used a couple of lectures to survey the topic content. I also gave a couple of lectures / practicums on searching for literature online and on reading articles. The second task was facilitating this rather complex arrangement.

I also provided requirements for how the groups should interact. They should bring a teaching handout -- a set of notes about the article that was written in such a way as to share with the partner. This handout could then form the basis of a fruitful discussion to understand the article. Similarly, when one pair taught the other pair, they should have a written handout of teaching material. The goal was to learn and discuss with colleagues, not chat about the topic. These handouts, as well as reflections, formed the basis of their group

notebook. The group notebook, particularly for the first three projects, was the primary assessment. If they were reading the articles, making notes to share, and actively reflecting upon the content, then they earned strong marks for the project. While the group notebooks were collated, each person was assessed individually for their contribution. (Academic Development Group 2006; Jones & Jones 2008).

2.3 Project 2: Word Learning

Project 2 was intended to be one step away from the very structured project 1 towards the entirely student-driven Project 4. The primary difference in organization of Project 2 was that, rather than my giving the list of sub-questions on the word-learning topic, the students would submit something that they were interested in regarding word learning, and then I would place them with others interested in something similar to form an expert group. The second difference between the two projects was that, rather than giving a group presentation at the end of the project, each student would write a short literature review that lead to a research question.

These seemingly small changes introduced significant confusion, however. When breaking it all down into steps, confusion is understandable. (1) I give an overall word learning topic question (which is a set up towards the final exam). (2) Each student submits a possible sub-question which I used to form expert groups. (3) With the expert groups formed, the students then had to come up with their 'real' sub-question together to resolve differences. Some groups got stuck here for multiple sessions and I eventually negotiated a question for the group. (4) Finally, each individual proposes a research question in their essay. It is easy to see how question after question could confuse.

During this project, I again led several sessions, including one practicum on experimental design and two overall lectures on word learning. Again, the students kept a project notebook of their reflections on the content and their handouts for discussion and teaching.

2.4 Project 3. Phrasal syntax

Project 3 was the next step towards total freedom on how they would conduct research in Project 4. While the focus in Projects 1 and 2 was on bringing in primary literature and teaching others, the focus on project 3 was on looking at existing knowledge, formulating a novel research question, and proposing a methodology to test that question. Therefore, the home groups and expert groups, with pairs and square arrangements, were gone. Instead people were

simply placed together and their job was to be a sounding board for other group members' research ideas.

I lectured several times during this brief project, with one lecture at the start on common child language research methodologies and three lectures on the acquisition of syntax, based out of the textbook. After this series of lectures, the groups met a couple of times to present their research ideas to their group and get help on designing the methodology. The individual kept a brief notebook of these interactions and their reflections on the topic. They also submitted an essay containing their research question and a method.

2.5 Project 4: Diagnostic Test for Dyslexia

In Project 4, students worked in groups to come up with a method of identifying children at risk for dyslexia. Similar to Project 3, they worked with only one group. There were only two requirements: (1) each student would turn in their proposal. (2) The group would document a plan for how they would research the issue and follow that plan. Assessment of this last group notebook would have no additional criteria from me. If they followed their own research plan, they would receive full marks. I lectured the first day of the project to introduce them to dyslexia and some contemporary researchers. I did not lecture again for the two and a half week period of the project.

The hope was that students would combine their practice working in groups from projects 1 and 2 with their practice forming research questions and methods from Project 3 to successfully complete Project 4. If they could do so, with very little present involvement from me, then it would be an overall successful semester.

Table 1: Overview of Course Designs Across Years

WEEK	2012 COURSE		2014 COURSE	
	CONTENT	ASSESSMENT	TASKS	PROJECT
1	Clark Ch. 1; 2-hour lecture		Clark Chs. 1&2; Lecturer gives topic question and sub-questions	Intro
2	Clark Ch. 2; 2-hour lecture		Clark Chapter 3; form home groups and expert groups; student experts do literature review	Speech
3	Clark Ch. 3; 2-hour lecture		Teaching of articles within expert group using pairs and squares	Speech

WEEK	2012 COURSE		2014 COURSE	
	CONTENT	ASSESSMENT	TASKS	PROJECT
4	Clark Ch. 4; 2-hour lecture	Annotated Bibliography due before Presentation	Integrate expert information in home groups; home group presentation	Speech
5	Clark Ch. 5; 1-hour lecture	Dyslexia Group 1 Presentation	Lecturer gives topic question; Clark, Chapters 4 & 6; students create sub-questions themselves; start literature review	Word learning
6	Clark Ch. 6; 1-hour lecture	Dyslexia Group 2 Presentation	Teaching others in expert group with shared sub-question learning	Word
7	Clark Ch. 7; 1-hour lecture	Dyslexia Group 3 Presentation; Take-Home Test 1 Due	Integrate expert information in home groups and develop a research question	Word learning
8	Clark Ch. 8; 1-hour lecture	Dyslexia Group 4 Presentation	Lecturer gives topic question; read Clark, chapters 7, 9 and 10. Students choose area of syntax and are placed in groups	Syntax
9	Clark Ch. 9; 1-hour lecture	Dyslexia Group 5 Presentation	Discuss research question and methods in group	Syntax
10	Clark Ch. 10; 1-hour lecture	Dyslexia Group 6 Presentation	Write essay with research question and methods	Syntax
11	Clark Ch. 11; 1-hour lecture	Dyslexia Group 7 Presentation; Take-Home Test 2	Students placed in groups; they decide all activities	Dyslexia
12	Clark Ch. 12; 1-hour lecture	Dyslexia Project Due	Group-designed activities	Dyslexia
13	Clark Ch. 14; 1-hour lecture		Dyslexia research question and methods essay due.	Wrap-Up

2.6 Course Evaluation Survey

An end-of-semester course evaluation was conducted. As part of a research project, ethics approval was obtained, and students had to provide consent for their responses to be used. The questionnaire consisted of a collection of Likert items (scale 1 to 5) and open-ended questions. The survey asked for students to rate:

1. Their view of the rate of absenteeism in the current paper relative to their other papers.

2. Interest in the topic at the start of the semester.

3. Interest in the topic at the end of the semester.

4. The number of lectures, as well as an open-ended question on why more, if more are desired.

5. Whether the group work was successful.

6. If it was not successful, an open-ended question asked why.

7. Whether the group work was enjoyable.

The evaluation ended with a series of open-ended questions including:

1. Did you contribute to the course as you wished?

2. Should the course be continued with this structure?

3. Should other courses follow this model?

4. Do you feel you learned more in this paper than other 300-level papers?

5. A space asking for any other comments.

The paper had been redesigned to increase engagement through a collaborative learning curriculum, and so many questions were targeted towards measuring those features. Other questions requested that the student make comparisons of the collaborative method against other courses that did not use such a paradigm. The results of this evaluation will be discussed as related points arise.

3. Evaluation of the Course

Evaluation of the course will be based upon four items. (1) Comparing marks and quality of work across the two years that the course has been run, (2) my own experience as the instructor of both classes, (3) observations of group interactions and (4) student feedback on the course.

3.1 Quality of Work across Years

The same essay questions that were used in 2012 for take-home tests were used again in 2014 for the final exam. Students were not aware of this fact during the semester. In 2012, students were given the questions and had to turn in a paper one week later. In 2014, students were given the questions one week before the final exam. During the exam, they could take all their notes and texts into the session, but had to write the essays during this time. Overall, the 2014 situation should have been more stressful as there was only a single 2-hour session in which they could write the essays. The same rubric was used for essay marking in both years.

In 2012, the overall mark on take-home test 1 was 79 and the mark on take-home test 2 was 84. While the 2014 exam did draw questions from both tests, very few students in 2012 actually answered the particular question from test 2, so it is difficult to compare that test with the 2014 final exam. The average mark on the exam in 2014 was an 81. Qualitatively, most of what students stated in answering the questions was accurate. Large-scale misunderstanding was uncommon. Where students lost points was in leaving out part of the required answer. Similarly, the average mark for the dyslexia project was 80 in 2012, but 83 in 2014. This was the case despite the fact that the 2014 students had a little over two weeks to complete the project, while the 2012 students had 8 weeks, where some work was required through group presentations, repeated dedication of time in lecture to the project, and an annotated bibliography.

In sum, all evidence points to the conclusion that students in the 2014 paper learned the linguistic content of the course at least as well as the students in the previous, more traditional format. The primary goal of the redesign was to focus on the larger take-away skills of self-designed learning and group work, so that if linguistic learning did not decrease and broader skills increased, then the design can be said to have met its purposes.

3.2 Student Perception of Learning

There is no direct way to compare learning in this course versus a broad selection of 300-level lecture-based courses. However, the survey did ask students what they believed their comparative learning to be. Of the 18 responses to this question, four believed they learned more or the same in the other lecture-based advanced courses they had taken. Comments included

☐ Because I'm more used to traditional format, I get more out of them

☐ About the same and not sure the process really helped

☐ Both. I keep all notes, textbooks, etc.

☐ Not sure compared to a religion paper, but I liked it [the current paper] better.

The other 14 of 18 responses believed that the current method was somehow more rewarding to them. Some of these responses focused upon skills. Such comments include:

☐ Skills, yes; information about the same

☐ I think I've learned more about people

☐ Since we did lots of researching, we know what to do if we want to know more about.

A number explicitly stated that content and skills from the current course would stick longer. Such comments include:

☐ I do think this paper will affect me more in the long run

☐ I always forget lecture material once the semester is done, but I feel that this class enabled me to engage with the material more so I will remember it more

☐ Yes, I learned way more in this paper and I def. will remember more of this. To be honest, I don't even remember anything from the other paper.

A few students directly tied possible better outcomes to their learning style:

☐ Will learn more; interacting with other people to share and explain ideas requires you to have an understanding of it

☐ I learn through interaction and doing rather than listening to a lecture or reading. In this way I found I learnt a lot more

☐ I feel I learned more from this one and would remember this one more as I actually had to really apply the knowledge to real-life situations (children) and design experiments.

Finally, one student stated, 'I think I've learnt skills more useful for further study in this class. I think this style does make the knowledge more my own & therefore I'd be more likely to remember it.' Just such an outcome was the goal of the course redesign. Most detailed content will be lost over time, and so it is a set of cognitive abilities that the student is gaining from a linguistics major.

Students were asked directly if other courses should follow this model and if the current course should continue along these lines. The most common answer was a qualified yes. Continue, but make fixes. Possible changes included more lectures, less shuffling of groups and a better match between assessment percentages and work.

The current version of the paper gave very little weight to the output of the early projects. If the student was actively participating in the group, represented by the group notebook, then the early projects would have very little long-term impact on their final mark. The hope was to decrease stress on the student until they've had some weeks to gain requisite skills. The flip-side of a low-cost assessment is what some students felt was a low-benefit assessment: significant work that does not strongly affect a result.

On top of the 'yes, but' answers were simple yes answers with reasons that matched the hopes of the paper. These students believed the current course offered interactive learning, chances to engage, development of skills for other areas, including confidence for the future and jobs, and was more realistic towards a work situation. One student did declare, 'Yes! This has literally been the best paper ever!' Regrettably, only one student was this effusive with praise for the course design.

3.3 Student Experience in the Course

Did the course design successfully encourage interaction, engagement and student ownership of their own learning? There are several ways to assess this, and the answer is mixed. As already discussed, several students specifically mentioned interaction as a key part of the course.

Additionally, the survey asked students their interest level at the start of the course and their interest level at the end, with a 1 being low interest and 5, high. The rating changed from 4.1 to 4.4. The lowest rating for any student for the end was 'neutral' (2 students). There was a strong ceiling effect with 55% of the respondents indicating they started at the highest level, so that no improvement was possible. In a pair-wise comparison, 8 students moved up over the course of the semester and 4 moved down. In sum, the course appears to have had a neutral to positive effect on student interest.

Another way to assess engagement is simple presence in class. It is not traditional at the university to keep attendance at lectures, a practice which I followed. However, I would estimate that on average 25% of the students would be absent on any given day. This could increase to almost 50% on the worst days. In a class based upon each student bringing their own contribution, this can be debilitating. On multiple occasions, only a single individual in an expert group of four would appear.

The major question is whether the class design contributed to this situation. Are the students specifically skipping this class because of its design? There are no absolute numbers with which to assess, but my impression is that the absenteeism was typical of classes at the university. I asked colleagues in many other programmes, and they all reported similar or worse numbers. I also asked the students in the class survey if they thought absenteeism was higher for this paper than others. Using a 5 point Likert scale with 5 being better than other classes and 1 being worse, the average guess from students was 3.75, and only a single individual said it was worse. Part of the purpose for the course redesign, however, was to increase student participation. The hope was that if students were actively engaged -- asking their own research questions, socially committing to a classmate to contribute – then they would be more present than normal. This *may* have occurred. A student knows they are responsible for something and so attends a class they might have skipped if it was a relatively passive lecture. On the other hand, it equally well might have occurred that the student had not yet done what they were intending to do, and so they chose to skip, rather than show up empty-handed. In short, this group project-based class had typical attendance for papers at the university.

The major factor, however, in determining the student experience in the class was the particular groups students found themselves in at any given time. Students rated group work success at 3.7 with the most popular rating of 4, and the enjoyment of group work at 3.4. When asked why group work was not successful at times, far and away, the answer was non-contributing group

members. Other team members did not bring anything to the class, did not respond to emails, or simply did not appear in class at all. Other reasons were given including shyness (3 responses), a noisy room (1 response) and lack of group direction (2 responses).

An additional question asked if the student themselves contributed the way that they wished. The most revealing answers indicated that they would contribute less if the topic or group was less interesting. Therefore, a cycle could occur in which a group was not working well, so a student would start contributing less, so the group would not work as well….

It is difficult to overestimate the importance of group dynamics to a course such as this. As a group moderator for half of the class, several group patterns were noticed.

1. Some groups flourished and were a true joy to watch. The highly functioning groups could be seen clustered together, everyone trading ideas, listening, and appearing to enjoy the process. Such engagement is extremely hard to generate in the best of lectures. If such groups were always the norm, then it would be easy to declare the current course design a true winner.

2. Another type of group was the one where the members were mostly present, but not very engaged with one another. One team member might sit on their computer typing, another looking over notes, and periodically, they would exchange an idea. This seemed to be a combination of personality type and preparation.

3. A very different issue could occur with what I will call the Excelling student. If one student in a group was operating at a higher level than the other students in their group, then the feedback for the excelling student was limited. There were instances in which a student would bring quite a set of sophisticated ideas, but the other group members were unable to effectively give feedback. This could only feel like a disappointment for such a student.

4. The final group pattern I noticed were the 'we've already done it' group. That group were working together, but they would have exchanged an idea or notes through email before the class session and therefore feel there was nothing more to discuss. Sometimes, such a group settled too quickly on something plausible, rather than truly striving for the best solution.

How can an instructor or a course design maximise the potential for the first time of group? There are several possibilities. First, even though the point is to increase independence, quick action from the instructor could assist. If a group is stuck on one matter due to simple disagreement, step in and negotiate a solution before multiple sessions are lost. Similarly, if it becomes clear that one group member is being abandoned by non-contributing members, move them to another group quickly. The way groups are put together can alleviate some issues. In the current paper, most early groups were randomly put together, except that expert groups were based upon interest. The final project groups, however, were tuned by myself to handle issues encountered during the semester. Criteria for the last groups included putting previously successful groups together again, connecting at least two Excelling students together, and distributing the frequently absent students around the groups. Many such adjustments can only be done after the course has progressed.

One further matter must be mentioned before leaving this topic. There are potential face threats (Brown & Levinson 1987) in all classroom interaction. However, in a class such as the one being described here, those potential face threats increase. The student is asked to frequently put themselves forth, presenting in writing and in verbal interaction their understanding of some difficult topic they barely know. These opinions can then be criticised. An argument can be made that such issues will be commonplace in work situations later so that getting practise negotiating face threats in a comparatively safe classroom environment is a good thing, but it makes them no less real. Time spent on developing team skills is rarely time wasted.

4. Final Evaluation

The goal of the course redesign was to increase a student's ability to learn independently by having repeated responsibilities towards other students. This groupwork-based project course does not work miracles. For a good number of students, the interactive, research-focused design appears to have been very rewarding. The understanding of content did not dip, and it may have improved. The argument for improvement is based upon similar results to the previous year under more difficult conditions. While the 2012 class had over eight weeks to do the dyslexia project, the 2014 class only had a little over two weeks, yet their proposals were of similar quality. Similarly, the 2012 class were able to spend as many hours as they wished over a week working

on an essay. The 2014 class produced just as high quality essays on the same topics in the more stressful examination environment. Moreover, several of the students did indeed seem to develop some of the independence skills and group work skills that were the primary goal of the course. This is most clearly seen in successful completion of the dyslexia project where I provided very little guidance. This was corroborated by the student survey. After all, after university education has ended, the odds of needing to know some fact about first language acquisition is rather low, but there is a significant chance that a former student will need to synthesise a great deal of material quickly, make choices about what actions to take next, teach others about what they have learned and argue for the next steps. This is what the course is intended to develop.

Even if this is all accepted, however, it does not mean that every course should function this way. Within a curriculum, if one or two courses had an organisation such as this, that would likely be sufficient for developing the specific skills being emphasised here. Other courses can emphasise other skills in a university's graduate profile. While many students found this pedagogical method rewarding, not all did, and there is no reason to force all courses to work this way.

The different components of the current course design need not all be used in future courses. The current paper had duelling aims. One aim, most clearly seen in project 1, was to use teaching others students to learn. This worked quite well. A second aim was to develop skills designing research, most clearly seen in project 4 on dyslexia. This also worked quite well. The transition between these two projects, however, was much messier. It is possible that one might run a course like project 1 throughout the semester; there is no requirement that it be paired with research design.

My take-home point, however, from this course was not related specifically to groups or projects. Rather, when the course was really working, it was because the student had a question, tried to answer it, had someone listen to them and then got substantial feedback. This is the critical moment. This *can* occur with groups, but it is not the groups that are fundamental; it is the feedback from another interested party. The current course is only one way to allow this to happen.

5. Epilogue

I am writing this article in 2015, one year after LING 318. Three of the students who were in LING 318 are now Honours students in a new class with me. The major task is for the three students and myself to conduct a research project together and write a joint paper. For this, we would jointly create the rubric for assessing this group project. The requirements of the rubric were that it should measure both individual and group contributions, and it must be documented for evaluation by another university. When I met with them, the students said, 'Why don't we do something like we did in Child Language with a group notebook showing what each person has done in the research plan?' I agreed.

References

Academic Development Group. 2006. *Information Series 400/2: Assessing Group Work*. Downloaded from http://www.canterbury.ac.nz/academicservices/info_series.shtml on 28 March, 2015.

Applebee, A. N. 1996. *Curriculum as Conversation: Transforming Traditions of Teaching and Learning*. Chicago: University of Chicago.

Brown, Penelope and Steven Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Clark, Eve. 2009. *First Language Acquisition (Second Edition)*. Cambridge: Cambridge University Press.

Dörnyei, Zoltán. 1997. 'Psychological processes in cooperative language learning: Group dynamics and motivation.' *The Modern Language Journal*, *81*: 482–493.

Barnett, Ronald and Kelly Coate. 2005. *Engaging the Curriculum in Higher Education*. Maidenhead, Berkshire, UK: Open University Press.

Conrad, Clifton and Laura Dunek. 2012. *Cultivating Inquiry Driven Learners: A College Education for the 21st Century*. Baltimore: Johns Hopkins Press.

Entwistle, Noel. 2009. *Teaching for Understanding at University: Deep Approaches and Distinctive Ways of Thinking*. London: Palgrave Macmillan.

Hughes, H. and A. Townley. 1994. *Cooperative Learning in Graduate Education: A Study of its Effectiveness in Administrator Training in two California Universities*. ERIC document ED 375517.

Johnson, David W., Roger T. Johnson and Karl A. Smith. 1998. 'Cooperative learning returns to college: What evidence is there that it works?' *Change, July/August*: 27–35.

Jones, Karrie A. And Jennifer L. Jones. 2008. 'Making cooperative learning work in the college classroom: An application of the 'Five Pillars' of cooperative learning to post-secondary instruction.' *The Journal of Effective Teaching, 8*(2): 61–76.

Spronken-Smith, Rachel and Tony Harland. 2009. 'Learning to teach with problem-based learning.' *Active Learning in Higher Education, 10*(138): 138–153.

Tagg, John. 2003. *The Learning Paradigm College*. San Francisco: Anker Publishing.

Whitehead, Alfred North. 1929. *The Aims of Education and Other Essays*. New York: The Free Press.

FIRST YEAR SOCIOLINGUISTICS AND THE TEACHING/RESEARCH NEXUS

Kevin Watson: *Department of Linguistics, University of Canterbury* *<kevin.watson@canterbury.ac.nz>*

Abstract

Although teaching and research have a symbiotic relationship in many institutions' mission statements, as academic staff we can sometimes view them as being direct competitors for our time, particularly perhaps with teaching at undergraduate level. Nevertheless, a considerable amount of our teaching at this level is at least *research-led*, where students learn about research findings in the field and also about the research interests of lecturing staff. It is arguably more difficult to follow a *research-based* approach, where students learn as researchers from a curriculum driven by inquiry-based activities. Hattie & Marsh (1996: 534) recommend that the teaching/research nexus should be driven by (1) the construction of knowledge by students rather than the imparting of knowledge by instructors, (2) the construction of assignments that reward deep learning, and (3) approaches which emphasize the uncertainty of the task. In this paper, I outline and reflect on several ways in which I have tried to engage undergraduate students in research activity, with the aim of having them produce original research projects. The main focus is an introductory sociolinguistics course in which I ask students to carry out a study of lexical variation in New Zealand by building upon Bauer and Bauer's (2003) 'Playground Talk' project. I discuss the benefits and outline several challenges of engaging first year undergraduate students in sociolinguistic research, and conclude that doing so is advantageous to the students, who can engage more effectively in the subject matter, and that it is also helpful for academics as we try to balance teaching and research commitments.

1. Introduction

Teaching and research have a symbiotic relationship in many institutions' mission statements, and as academics we are often encouraged to reflect on how our work connects with the teaching/research 'nexus'. Typically, however, the practices and discourses surrounding teaching and research reflect separation rather than symbiosis. At the institutional level we might be recognised and rewarded differently for each activity (e.g. in promotion applications, where conventional wisdom is that success in research is viewed more favourably than success in teaching), and at the systemic level we must apply to different funding bodies for research and teaching (e.g. in New Zealand, the Marsden Fund for research vs Ako Aotearoa for teaching). At the personal level, we might simply see teaching and research as direct competitors for our time. Nevertheless, many academics are expected to do both of these activities, occupying what Tennant et al. (2009: 170) call a 'hybrid', often contradictory, space. In this paper, I outline and reflect on how I have tried to engage students in research activity. I focus on a first year undergraduate course, because it is arguably at this level where the tension between research and teaching is most keenly felt (Zamorski 2002). I begin with a discussion of some of the scholarly work which has focussed on the teaching/research nexus, including how the relationship has been conceived across different disciplines and academic levels. After this I introduce the first year course on which I focus for the remainder of the paper, discussing first its general structure and main content, before elaborating on the assessment tasks which are designed to develop the students' research skills. I then discuss the benefits of such an approach, along with the challenges, before concluding by outlining the lessons I have learned by teaching the course in this way.

2. The teaching/research nexus

The relationship between teaching and research is often regarded as a key facet of a university, but this was not always the case. As an example from the New Zealand context, Robertson & Bond (2005) outline the changing perspectives on the teaching/research nexus over time, using the history of the University of Canterbury (UC) as a case study. They note four 'phases' in UC's history, and chart the changing relationship between teaching and research. The first phase (1870-1945) includes the earliest days of Canterbury College when,

Robertson & Bond (2005: 514) write, 'Educationally, research was not of interest. Teaching, with a concern for mastering a body of knowledge, was emphasised'. The second phase (1945-1946) is described as a 'watershed' for the relationship between teaching and research, not least because of the publication of a manifesto instigated by Karl Popper (Allan et al. 1945) which stated that research and teaching should be seen 'not as separate functions of a University teacher but as complimentary parts of a single entity' (Allan et al. 1945: 2, cited in Robertson & Bond 2005: 519). Although this call began to be heeded in spirit, in practice it was taken as the addition of research to the existing suite of activities being conducted: 'research was to be encouraged but teaching and research were seen as separate functions' (Robertson & Bond 2005: 522). Research and teaching are reported to be more closely linked in phases 3 (1946-1990) and 4 (post 1990), by which time the teaching/research nexus had become embedded into UC's charter. It continues to be positioned prominently in the University's planning documents. For example, UC's Learning and Teaching Plan 2013-2017 states that the university 'has at its very core research-led learning and teaching' (p.2) and that 'a key aim is to encourage staff to identify ways in which their teaching can be improved by their research activities' (p.2). Similar aims can be found in the documentation of many other institutions, in New Zealand and elsewhere. But what is the reality behind such aims? What, in practice, is the relationship between teaching and research in the day-to-day work of academics? In a seminal meta-analysis addressing this question, Hattie and Marsh (1996), argue that there is in fact no overall relationship between research and teaching. They reviewed 58 studies on the matter and quantified various factors (e.g. those related to research, such as the number of publications or citations vs. those related to teaching, such as student evaluations, as well as others such as different discipline areas), and they then tested for correlations between the factors as a way of assessing the relationship between research and teaching. In total, Hattie and Marsh (1996) tested 498 correlations, but found that only 20% of them were statistically significant. On the basis of this they concluded that 'the common belief that research and teaching are inextricably entwined is an enduring myth' (p. 529). The idea that the teaching/research nexus is based on mythology is put forward elsewhere as well (e.g. Hughes 2005), but should be treated cautiously when based on quantitative measures alone. It is not clear that we should expect there to be a relationship between, for example, research *productivity*, measured by number of outputs, and teaching *quality*, measured by student evaluations, but this is exactly what Hattie and Marsh

(1996) imply when they agree with Feldman (1987: 275) who writes that 'the likelihood that research productivity actually benefits teaching is extremely small… Productivity in research and scholarship does not seem to detract from being an effective teacher'.

When work on the teaching/research nexus has used a qualitative methodology, where, for example, university staff (e.g. Newmann 1993) and students (Neumann 1994, Robertson and Beckler 2006) are interviewed about their understandings of the relationship, the conclusions are typically more positive than the results that have been presented in quantitative studies. Neumann (1994: 326), for instance, finds that students of all academic levels and in all disciplinary groups 'discussed subjects which they had enjoyed because the lecturer was at the forefront of knowledge, relevant examples from the teacher's research were used in teaching, and students were taught useful techniques which the lecturer used in his or her own research'. Neumann (1994: 326) also comments that this view tends to manifest more strongly at more advanced student levels, perhaps unsurprisingly, and particularly in the sciences and social sciences (vs. e.g. humanities). Science students also commented positively on being able to 'do experiments…or were able to undertake research projects in close contact with a lecturer where they realised for the first time that in science "things don't always work out"' (Neumann 1994: 327). These are examples of what Newmann (1993) calls the *tangible* nexus between teaching and research, where we see a focus on the transmission of knowledge and skills. This contrasts with the *intangible* nexus, which concerns the transmission of approaches and attitudes to knowledge, particularly 'imparting to students a questioning, critical approach to knowledge, as well as a positive attitude to learning' (Neumann 1994: 327). One way in which these attributes are developed, Newmann (1994) suggests, is through the use of assessments which encourage students to engage in research of some sort, for example, devising an experiment, adapting a case study in some way, or undertaking a small research project. Newmann (1994: 330) describes students' typical responses to these assessments: 'many of the students described their surprise, firstly at being asked to do something different from the usual assignments and secondly their sense of fear – sometimes coupled with excitement – at the challenge. All stated that by the time they had completed the assignment they had found the work intellectually stimulating and enjoyable.' There is a real sense, then, that students themselves believe the teaching/research nexus to be important. For them, positive things

happen when teaching and research are meaningfully combined, even if quantitative measures of the relationship suggest otherwise.

2.1 The teaching/research nexus across disciplines

Robertson & Blackler (2006) focus on students' understanding of research in three different disciplines: physics, geography and English. For physics students, research is 'esoteric' and 'remote' from them, often done in 'another language'. It is something that is done 'higher up', by lecturers, not undergraduates. For geography students, research is 'coming up with and finding an answer to a question' and is something they felt able to engage in because of a 'shared sense of research community and methodology'. For English students, research is more personal, and although there is a shared sense of community there is also the view that lecturers do research 'at a different level'. Unlike in physics, there was the sense that English students are at least 'speaking the same language' so that even first year students felt able to engage with research activity'. Robertson & Blackler (2006: 224-5) argue that these differences are due to differences in disciplinary knowledge structure. For example, in physics 'knowledge tends to be cumulative…and the relationship between research, teaching and learning hierarchical', whereas in English 'the "flatter" more accessible structure of knowledge means that it is possible for students to engage with that knowledge earlier and through their own research'. While we should be cautious about seeing very firm and fixed boundaries between these disciplines – work in English obviously can be and often is cumulative – it is worth taking a moment to consider how linguistics would fit and, perhaps, even whether it is possible to categorise linguistics as a discipline in these terms. Some linguistics certainly requires considerable cumulative knowledge. It is highly unlikely, for example, that any first year student of linguistics would be able to engage with the complex notions and impenetrable-to-the-outsider formalisms in papers such as 'A minimalist condition on semantic reconstruction' (Ruys 2015), or indeed many of the other papers in *Linguistic Inquiry*. In this sense, linguistics is similar to the work in physics which students see as being in 'another language'. But other sorts of linguistics, such as the paper 'Ideologies of language and race in US media discourse about the Trayvon Martin shooting' (Hodges, 2015) and many others in *Language in Society*, are likely to be more straightforward for beginning students to engage with, even if we might not expect them to fully understand all of the theoretical nuances. Differences such as these make it

unreasonable to expect the teaching/research nexus to be the same for teachers in each of these (sub)fields.

2.2 The teaching/research nexus across levels

As well as the (sub-)disciplinary differences, there are also differences in the relationship between teaching and research across different academic levels. Of course we expect PhD and Masters students to be involved in research, and also, in the New Zealand context, students doing their Honours year (typically in the year following their Bachelor degree) are required to carry out their own research project. At undergraduate level there is more variation both in terms of how and when students are engaged in research, and precisely which kind of research they do. It is common for undergraduate lectures to be research-led, where students learn about research findings in the field and also about the research interests of lecturing staff. This is probably standard practice in all Linguistics programmes. It is arguably more difficult at undergraduate level to follow a *research-based* approach, where students learn as researchers from a curriculum driven by inquiry-based activities. Indeed, Zamorksi (2002: 417) reports that undergraduates often believe themselves to be the 'recipients of research, rather than actors in its production'. The perspective of the lecturers who Zamorski (2002) interviewed was that teaching had to be research-led in the early part of a programme, because it was important the students were taught the relevant content before they embarked on research. Zamorski (2002: 422) notes how this view is that research fits 'into a linear intellectual maturation process, from intellectual dependence to intellectual independence', contrasting this with the view of students, who say that even at the end of an undergraduate degree programme 'whatever research experience or work undergraduates undertook, it was not usually regarded as "proper" research' (2002: 419).

There are a wide range of reasons for undergraduates' views that they do not do 'real research'. One is that even in research-led programmes, undergraduates rarely see research as a process. We typically focus on delivering the product – the results and theoretical implications – not the details of what happened, how and why, or of what went wrong. Presenting only the final, public face of a research paper can give the impression that it is only student research which is 'messy', and so less 'real' than published work. Developing students' awareness of the processes through which knowledge is created and constructed is an important step to developing a 'culture of

inquiry' (Robertson & Blackler 2006: 215), something which needs to be 'embedded in the curriculum from day one if undergraduates are to understand the value of their own research' (Garde-Hansen & Calvert 2007: 114). Giving a behind-the-scenes account of a research paper in an undergraduate class is different from simply teaching research methodology, and it is only really possible when we teach about research that we have a detailed, often personal, knowledge of. It is perhaps in this way that our research and our teaching can be clearly linked, by researchers bringing their personal research experiences into the classroom.

Hattie and Marsh (1996: 534) make a number of other recommendations for how the relationship between teaching and research can be strengthened, including (1) emphasising the construction of knowledge by students rather than the imparting of knowledge by instructors, (2) constructing assignments that reward deep rather than surface learning, (3) developing strategies across all disciplines that emphasize the uncertainty of the task, and (4) ensuring that students experience the process of artistic and scientific productivity. In the remainder of this paper, while building on the discussion above, I discuss how these recommendations can be usefully incorporated into a first year sociolinguistics course.

3. Introduction to LING102

Two 100 level courses are required in order to major in Linguistics at the University of Canterbury: LING101 and LING102. LING101 is an introduction to English phonetics, phonology, morphology and syntax, and LING102 is an introduction to sociolinguistics. Typically, LING101 runs in Semester 1, and LING102 runs in Semester 2. Students majoring in Linguistics normally take LING101 before LING102 but this is not required, as LING101 is not a pre-requisite for LING102. Students are allowed to take LING102 before LING101, and non-majors might take only one of these two courses. This means that LING102 has to be designed to work both for students who have taken LING101, and have therefore been told about phonemes, allophones, morphemes, parts of speech and phrase and clause structure, among other things, and also those who have not taken LING101, and who are often coming to the discipline of Linguistics entirely from scratch, with no knowledge of key concepts or terminology. Furthermore, many students take

LING102 as an interest paper, and do not follow it with other higher level linguistics courses. For some, then, LING102 is the beginning and the end of their experience of linguistics.

As with any introductory course, it is not possible in LING102 to cover every aspect of the vast discipline of sociolinguistics. Topics are selected according to the requirements of higher level courses the students might take in our department in future. For example, we have higher level courses in variationist sociolinguistics, but not currently in discourse analysis, so the content of LING102 is geared towards the former, and the students are given only a taste of some main themes and methodologies of the latter. An important and obvious objective of the course, then, is to provide relevant training for students, such as our linguistics majors, who will do further higher level work in linguistics. But this cannot be the only objective because, as noted above, some students' experience of linguistics stops here. What do these students gain from the course? The same question can be asked of linguistics majors. Even if LING102 provides useful training for higher level linguistics courses, what is its long term value? As Hatfield (2014) notes, most linguistics majors will not become professional linguists. Most will not do any linguistics of any kind after their undergraduate degree. The value of the course comes in at least two forms. First, it is hoped that, in line with all other similar courses in sociolinguistics, LING102 has some long term value because of the social relevance of its content (in this case, tackling language myths and prescriptive vs descriptive views, other aspects of social and linguistic stereotyping, etc – see below). Second, skills needed to be able to carry out research – critical thinking, engaging with literature, data coding, quantification techniques, IT skills – are embedded throughout the course, in the hope that they are useful not only for linguistics majors doing higher level linguistics work, but also for students doing higher level work in any major and, since they are largely highly transferrable skills, also more broadly. The aim is that students do not just learn about sociolinguistic research, but that they also carry out their own original work, learning the necessary skills along the way.

Before I describe the assessment in the course, I first provide some background information about its structure and the general topics covered. There are 3 contact hours per week: two 1-hour lectures and one 1-hour tutorial. The lectures follow more or less a standard structure, with a single lecturer talking to a group of around 80 students in a tiered lecture theatre, with some limited opportunities for discussion. Lectures are the main way in

which new subject content is delivered to students. The content of the course is divided into five 'blocks', each covering a different theme in sociolinguistics, broadly conceived. The two largest blocks, in terms of the number of weeks dedicated to them, are: (1) 'What does our language say about us?', in which we explore Labov's key variationist sociolinguistic work, with follow up examples and case studies from work in a New Zealand context, and (2) 'Can language affect how we experience the world?', in which we focus on language attitudes, including linguistic stereotyping. In these blocks, and the others, alongside the main results and the broader implications, there is a considerable focus on research methods. The students are taught not only about what the results are and what they mean, but also how the research was carried out. Examples from the research projects at the University of Canterbury and the New Zealand Institute of Language, Brain and Behaviour give added authenticity to this work and offer opportunities for the sorts of behind-the-scenes accounts mention above. Methods are then the main focus in the weekly tutorials, where students are given the opportunity, for example, to explore datasets, create tables and graphs, and consider how they would describe and explain the results they see. As well as this, more general research skills are taught in tutorials, including how to carry out a literature search, how to read critically, and how to write a literature review.

4. The LING102 research report

The focus on research skills is important because the main piece of assessment in LING102 involves students carrying out a research project, which is worth 40% of the final grade. Students can choose from two (or, in some years, three) topics for their project, and I focus on one of them here. This involves extending the work presented in Bauer and Bauer's (2003) *Playground Talk* project. This project set out to investigate regional variation in New Zealand lexis, by focusing on playground vocabulary of children in primary and intermediate schools. In total, 150 schools from Kaitaia to Bluff participated (Bauer and Bauer 2003: 2), which involved teachers completing a paper questionnaire, after discussion with pupils, and answering questions about playground games and the sorts of words and phrases the children were likely to say in given situations. Two examples are:

(1) Brackie goes to the shop and buys a mixed bag of things like jaffas,

winegums, toffees, Minties, and liquorice all-sorts. How would you finish the sentence "Brackie's got a bag of _____"?[1]

(2) At your school, do children play a game with many players where one player has to run and try to touch another player while all the other players try to run away and not get touched? What is this game usually called at your school? At your school, how do you usually tell someone that they are to be the player who tries to touch the others? Is there a special word which the player who touches someone says as they touch them?

Many of the questions in Bauer and Bauer's (2003) project, along with some additions, were used to create an online questionnaire via *GoogleDocs*. The online questionnaire needed to be more wide-ranging than Bauer and Bauer's original so it had a somewhat more general range of questions, in 4 categories: *Words for Things* (which included question like 1 above), *Our School Days* (which included questions like 2 above), *Feelings and States* (which included questions like 'What word would you use to describe somebody who has become intoxicated from too many alcoholic drinks?'), and *Friends, Family and Being Social* (which included questions like 'What word would you use to describe an attractive person?'). As well as this, the intention was for the questionnaire to be completed by as many different people as possible, not just school children. Participants were therefore required to provide sociodemographic information, such as their age, sex, country and city of origin, and whether they have lived outside of the country/city of origin for longer than 3 months.

The results from the online questionnaire form the raw data on which the LING102 students base their research projects. The summary instructions given to the students are:

> For this topic, you will investigate how words are used in New Zealand. You will write a research report to explore this. The exact focus of the report is up to you. You may choose to answer questions such as: How have words in New Zealand changed over time? Are there regional differences between words across the country? Are there differences between the way women use certain words and the way men use them? You do not need to answer all of these questions but can choose a subset. You could in fact ask different questions entirely, as long as they can be answered with the data we generate. The direction

of this research report should be driven by your own thinking and reading around this subject.

The data will come from a survey of lexical variation in New Zealand. It is an online survey, which asks questions about words. Your task is first to spread the word about this survey, via your friends and in other ways (e.g. via social media). We need lots of people to complete the survey so we have some new data to work with. After a while, you will be given access to a set of the raw data. Your task for the research report is to analyse some of the data to answer your questions. You are expected to connect the results to the literature we have discussed in class, and to additional reading you have done in preparation for this research project.

This is a 'real life' research project. The data has not been manufactured specifically for this class. It has not been 'cleaned up'. It is new, it is being generated right now, and it will come from real people. Some people who fill in the survey will make mistakes, some will not take the task seriously, and we do not know what the results will be like. They may look messy at first, but it is our job as analysts to understand the dataset, and to seek out any patterns. When you have completed your project, you will have contributed to our wider understanding of lexical variation in New Zealand.

The first task for the students is to share the link to the questionnaire among their friends and contacts, for example via social media. This happens early in the course, to give enough time for some results to be generated. The students are told that they are required to write a research report about lexical variation, but within that broad objective there is considerable flexibility. For example, I do not set a particular research question which must be answered, but instead offer a range of options and also tell students that they can make other suggestions, if they wish (see e.g. in the first paragraph of the instructions above). I also give them the flexibility to decide which questionnaire questions they focus on (e.g. they can focus on one thematic category, such as *Our School Days*, or some other subset of questions). I explain that providing a suitable rationale for these sorts of decisions is part of the research process.

At the time of writing we have responses to the questionnaire from 3000 informants, but this is far more than is necessary for or manageable in a first year class. Students are therefore provided with the raw data from about 500

informants. They must assess the usefulness of the dataset for answering their chosen research question(s). On one occasion, for example, a student was interested in comparing New Zealand English words with those from England, Scotland, the United States and Canada. The dataset I initially provided was not suitable for this line of enquiry, but an analysis was possible with a different subset from the larger sample, so after some discussion a new subset was generated which was more appropriately balanced.

Once in possession of the raw data, there are a number of steps the students need to carry out before the data can be analysed. One important step is to standardise the spellings of entries which are misspelled, or which have quotation marks, or some other idiosyncrasy. This is important if the quantification of the results is to be automated, as the spreadsheet software would otherwise incorrectly treat e.g. *New Zealand*, *NewZealand* and *NZ* as different categories of response. Also, it may be necessary for students to recode the data so that new categories are available for comparison. For example, a student might wish to examine differences between the North and South Island of New Zealand, so the particular locality of the participant would need to be recoded into these binary categories. Students must also assess the equivalence of lexical items given in response to the questions, and decide whether, for example, in a question about swimwear, *costume*, *cozzie* and *cossie* are the same or different. The steps are important in both practical (i.e. the use of spreadsheet software in data coding) and theoretical (i.e. the equivalence of variants of a variable) terms. They are also important in ensuring students explore the dataset in depth, and that they experience the processes behind this sort of analysis.

The first analytical step, once the data has been appropriately recoded, is to quantify the results. This involves using the spreadsheet to generate token counts, typically with pivot tables to automate the process as much as possible. It will also usually involve the calculation of percentages. Both of these tasks, using a spreadsheet for automatic data quantification and for calculating percentages, sit firmly outside the comfort zone of many of my students at this level. Because of this, students are given a considerable amount of training in this area. This year for the first time these skills have been foregrounded in some of the tutorials, giving the students the opportunity to work in a computer lab, with the support of a tutor, developing their spreadsheet (and, perhaps, their numeracy) skills. As well as this, short videos are made available, focusing on, for example, how to make pivot tables and how to use formulae, so that the students have as much information available as possible.

And, finally, to firmly focus students' attention on the fact that these skills are important, I set a 'take home task', earlier in the course and worth 10% of the course grade, which requires them to engage with these sorts of analyses (in a much more constrained way, and on a different dataset from that which will be used for their main reports) many weeks before the deadline for the final research report.

This research task efficiently combines many of the recommendations Hattie and Marsh (1996) have for strengthening the teaching/research nexus, mentioned above. The emphasis is on the students' construction of knowledge, not on the transmission of knowledge from the lecturer. They know that I, like them, do not know in advance what the results will be. This builds in an element of uncertainty and risk. The students also know that they are not only contributing to their own knowledge, but they have the potential to advance our knowledge in the field, because they are doing 'real' research. This promotes deep rather than surface learning in a range of ways. For example, students must engage with the data, interact critically with the relevant theoretical literature, and apply some of the theories to their dataset. They cannot just learn the content of the lectures and reiterate it, nor can they simply read and summarise published research papers. These strategies might work reasonably well for an essay, but they would not be successful for this research report where the students generate new knowledge. As there has been much less sociolinguistic work on lexical variation than on, say, phonological or grammatical variation in New Zealand English, there is the opportunity for these first year undergraduates to make an impact. Indeed, the students who uncover the most interesting patterns each year are encouraged to collaborate and write up the results for publication. We are fortunate in New Zealand to have a suitable venue for interesting but smaller scale work – the *New Zealand English Journal* – and some work from this course has started to appear in this publication (Watson et al. 2013). This is a major contributing factor in trying to enhance the teaching/research nexus in my own practice, by 'dismantling the notion that that research into the discipline [is] authorative and precious, such that these students could take ownership of the research material' (Garde-Hansen & Calvert 2007: 109).

5. On the benefits and challenges

In this section I consider whether the objectives of the course were met, by reflecting on feedback from course evaluation surveys and additional comments received by students via email. The course is overwhelmingly well received. One person (out of a total of 72 who did feedback questionnaires) reported that he or she would prefer a longer essay than a research report, but otherwise the research task received positive feedback, even if sometimes with a caveat, e.g. "The research report wasn't even that bad. It was quite exciting to research something in that way"; "The research report gave me a chance to research things we weren't necessarily taught in class. I had to read wider than class information to answer my research question. Because I enjoyed what I was learning so I understood it easily and it didn't feel like work!" Many students cited the research task as being a main factor in their enjoyment of the course, and some commented that it had changed their opinion of the field: "I was not particularly interested in sociolinguistics before taking this course but I have come to understand why this is an exciting area of research"; "I now hope to take linguistics as my major in my degree, with an emphasis on sociolinguistics"; "I may have changed my career path now to Linguistics." One student sent me extended feedback in an email. I reproduce it below (with permission), as it connects to a number of issues I have touched on above:

> When we were first tasked with carrying out our own research assignment, I was a little scared because it seemed like such a big job and something that was totally new to me. The lecturer was very reassuring, however, and I soon realised that the sooner I started it, the better. I found the tutorials immensely helpful, I learned to use Excel in new ways to deal with the data. It was the most challenging assignment I was given in my 100 levels papers, but in the end it was very satisfying to accomplish it. I learnt that when you undertake a research task, you really don't know what results, if any, you will uncover by the end. I was very glad that I started as soon as the second half of semester began, as I think I would have really struggled with the workload if I had waited until just before it was due. I am hoping to study at a postgraduate level, and so I am really happy to have been thrown into this research project early in my university studies, I feel that it alleviated some fears around research moving forward.

The benefits of including the research aspect of this course are: (1) some students find it exciting, because they can engage with class material in

new ways, (2) some students are encouraged to read more widely around the topic, (3) it increases some students' engagement and affinity with the field of linguistics, (4) although daunting at the beginning for some students, completing the research project is highly satisfying, and this can alleviate possible future fears, and (5) it increases core transferrable skills (e.g. using spreadsheets in new ways).

As well as these benefits, there are practical challenges to teaching the course in this way, and many of them relate to the fact that it is a first year class. This means, firstly, that the class is reasonably large (around 80 students per year). This is by no means unmanageable for a typical lecture/tutorial course, but when students are all doing research projects, with a mix of different research questions, the staff time needed outside of class hours is considerably increased. This can be mitigated, by providing a wide range of resources online, including sets of 'how to' videos for aspects of data management, but this course still requires more resources than it would with more typical assessments of essays/examinations. The combination of the high numbers of students and the flexibility they have in choosing their own research questions also impacts on the time required for marking. Because the focus of each student's report could be slightly different, assessing the report is more challenging than when students are able to choose from only 5 or 6 essay questions. The marker must read the report itself and look in detail at the student's analysis of the data (which they are also required to submit in spreadsheet form), increasing the complexity of the marking task and the time required to do it.

Other challenges arise from first year students being 'first timers', in perhaps multiple ways. Many students are linguistics first timers, if, as is often the case, they have not taken LING101 before this course. This is by no means a major obstacle, because the course is designed to take it into account, but it does introduce a challenge when students want to come up with their own reading lists. Sometimes the work students want to read is too advanced for them, when they do not yet have even one full linguistics course behind them. This can often be remedied in office hours, but it is an additional cost on staff time. Many students are also statistics first timers. They are not required to carry out statistical testing in this course, because there is no time to teach them properly how to do it, but sometimes students are also not completely confident with basic numeracy skills. Often this lack of confidence is driven by fear rather than lack of ability, but this fear is something which must be

overcome if the students are to be able to cope with the simple quantification of the results in their research report. The tutorials help with this, as does the separate take-home task which gives them practice of the sort of numeracy skills they will need. Many students are also computer software first timers. Although it is tempting to think of most first year students as 'digital natives' (Prensky 2001), who are 'IT-savvy' and who can easily open and save files, and use Microsoft Office, this is an over-generalisation. Of course IT-savvy students exist, but we should not expect this to be the norm. Developing students' skills in this area is highly valuable, but again it is a cost on staff time. Finally, first year students are likely to be research first timers, who have likely never undertaken a research project before, and who, perhaps more importantly, may not come into the class expecting one. The research is ultimately rewarding, but is at first daunting, as the student feedback indicates, and this needs to be taken into account when explaining the assessment structure of the course. It also needs to be borne in mind continuously as the course progresses, as it is easy for fear to reappear and derail good student work.

6. Conclusion

I conclude with some brief comments on the lessons I have learned by teaching LING102 in this way. First, in order to be successful, the research task needs to be as much like a 'real research project' as possible, or it does not excite the students in the same way. This means that the project should generate some new data, and that the students should be involved in its collection. Both of these aspects introduce an element of risk, which is another key element in increasing the authenticity of the task. But also important is the need to exert a degree of control over the potential risks. If, in a typical research project, the data collection fails for some reason, or if the data is problematic or the results not conclusive, then this is likely a considerable disappointment but it is not completely disastrous because things can usually be redone. Within the constraints of a classroom, particularly on a first year course with a high number of students, this is much more difficult, so steps must be taken to ensure the data collection will yield suitable data, and that the results will likely lend themselves to being written up in a report. For LING102 this was achieved by having the students use data from a survey that I developed (based on Bauer and Bauer's (2003) solid foundation). Of course there are many benefits of having students create and execute their own data

collection methodology, but this was thought to be too much risk in this sort of introductory course.

A final lesson concerns the teaching/research nexus. Much of the research on this relationship examines the effect of research on teaching. That is, the surrounding discourse typically discusses approaches such as 'research-led teaching', or 'research-based teaching', examining the benefits of the teaching/research nexus for teaching. There is much less discussion of the benefits on research. But in the case of LING102, the research project aspect of the course becomes a shared research goal between students and staff. When we begin, nobody knows what patterns will emerge, and we learn as the course progresses through collaborative research practice (see Garde-Hansen & Calvert 2007: 108). The project encourages and facilitates my own thinking on the topic, and on related research. It is can therefore be important from a research perspective as well as from the perspective of pedagogy.

Note

1 'Brackie' is one of a number of invented characters Bauer and Bauer (2003) used so as not to inadvertently implicate real children in any of the scenarios that were presented. See http://www.victoria.ac.nz/lals/research/projects/language-in-the-playground-project/table-of-contents for further details of the project, including the full questionnaire. Thanks to Laurie Bauer for granting permission to use some of the questions for the LING102 project.

References

Allan, R., Eccles, J., Forder, H. Packer, J., Parton, H. and Popper, K. 1945. *Research and the University*. Christchurch: University of Canterbury.

Bauer, Laurie and Winifred Bauer. 2003. *Playground Talk*. Wellington: Victoria University, School of Linguistics and Applied Language Studies.

Jenkins, Alan, Mick Healey and Roger Zetter. 2007. *Linking Teaching and Research in Departments and Disciplines*. York: The Higher Education Academy.

Feldman, Kenneth, A. 1987. 'Research productivity and scholarly accomplishment of college teachers as related to their instructional effectiveness: A review and exploration'. *Research in Higher Education* 26: 227–298.

Griffiths, Ron. 2004. 'Knowledge production and the research teaching nexus: The case of the built environment disciplines'. *Studies in Higher Education* 29 (6): 709–726.

Garde-Hansen, Joanne and Ben Calvert. 2007. 'Developing a research culture in the undergraduate curriculum'. *Active Learning in Higher Education* 8(2): 105–116.

Hatfield, Hunter. 2014. 'Linguistics as education for the future non-linguist'. Paper presented at the *Teaching Linguistics to Non-Linguists* workshop, University of Canterbury, 7 July 2014.

Hattie, John and H. W. Marsh. 1996. 'The Relationship Between Research and Teaching: A Meta-Analysis'. *Review of Educational Research* 66(4): 507–542.

Hodges, Adam. 2015. 'Ideologies of language and race in US media discourse about the Trayvon Martin shooting'. *Language in Society* 44(3): 401–423.

Hughes, Mark. 2005. 'The mythology of research and teaching relationships in universities.' In: Barnett, Ronald (ed) *Reshaping the University: New Relationships Between Research, Scholarship and Teaching*. Buckingham: Open University Press. 14–26.

Newmann, Ruth. 1993. 'Research and Scholarship: Perceptions of Senior Academic Administrators'. *Higher Education* 25: 97–110.

Neumann, Ruth. 1994. 'The teaching-research nexus: applying a framework to university students' learning experiences'. *European Journal of Education* 29(3): 323–338.

Prensky, Marc. 2001. 'Digital Natives, Digital Immigrants'. *On the Horizon* 9(5): 1–6.

Robertson, Jane and Carol Bond. 2005. 'The research/teaching relation: a view from the 'edge''. *Higher Education* 50: 509–535.

Robertson, Jane and Gillian Blackler. 2006. 'Students' experiences of learning in a research environment'. *Higher Education Research & Development* 25(3): 215–229.

Ruys, E. G. 2015. 'A Minimalist Condition on Semantic Reconstruction'. *Linguistic Inquiry* 46(3): 453–488.

Tennant, Mark, Cathi McMullen and Dan Kaczynski. 2009. *Teaching, Learning and Research in Higher Education*. Florence: Taylor and Francis.

Watson, Kevin, Sarah van Eyndhoven, Alina Filatova and Ashley Harrison. 2013. 'Lexical variation in New Zealand English: an online apparent time study'. *New Zealand English Journal* 27: 33–45.

Zamorski, Barbara. 2002. 'Research-led teaching and learning in higher education: A case'. *Teaching in Higher Education* 7(4): 411–427.

THE EDITOR'S NOTICEBOARD

Kevin Watson: *Department of Linguistics, University of Canterbury*
<kevin.watson@canterbury.ac.nz>

I am pleased to bring you volume 58 of *Te Reo*, the third one published in 2015. I am especially pleased to be able to bring the journal back up to date. To do this, it was necessary for the reviewers to give efficient feedback, and for authors to action their recommendations quickly. Thanks to all for making it happen.

The second half of this volume is a special section consisting of papers from a workshop on teaching linguistics and sociolinguistics, which I organised at the University of Canterbury in April 2014. I thank UC's School of Language, Social and Political Sciences for financial support for the workshop. The workshop was a result of discussions during the Subject Meeting at the 2013 NZ Linguistic Society Conference, where it was thought that colleagues would likely find discussions of pedagogical practice interesting and valuable. Following this, in 2014 and 2015, there have been teaching streams during the NZLS conference meetings. *Te Reo* (2007) was a special issue on linguistics pedagogy, and the papers in the second half of this volume continue that theme. I encourage colleagues who have ideas for papers relating to linguistics and sociolinguistics pedagogy to get in touch with the editor.

The next volume, *Te Reo* 59, will, all being well, appear in its scheduled year – 2016. The current deadline for submission of papers to be published in *Te Reo* 59 is 1st April 2016. As always, please continue to consider *Te Reo* as a possible venue for publication of your work, and contact the editor if you have a proposal for a special issue or themed section.