Vocabulary profiles as predictors of the academic performance of Teaching English as a Second Language trainees

Lori Morris*, Tom Cobb

Département de linguistique et de didactique des langues, Université du Québec à Montréal,
C.P. 8888 Succursale Centre-Ville, Montréal, QC, Canada H3C 3P8

Received 28 October 2002; received in revised form 1 May 2003; accepted 7 May 2003

Abstract

Given the cost of running second or foreign language teacher training programs with a high number of in-school placement hours, the selection of candidates who are likely perform well academically and complete their studies is an important consideration. The purpose of the present study was to examine the potential offered by vocabulary profiles as predictors of academic performance in undergraduate Teaching English as a Second Language (TESL) programs. To this end, vocabulary profiles were established for 122 TESL students by means of an analysis of 300-word samples of their writing. The students’ scores on each profile component were then correlated with the grades they were awarded in two of the grammar courses in their program of study. Finally, the effect of the students’ mother tongue on both their vocabulary profiles and academic results was considered. The findings of the study reveal that the students’ vocabulary profile results correlated significantly with grades in the more procedurally oriented of the two courses. Furthermore, vocabulary profiles proved to be useful in carrying out a finer assessment of the language skills of high proficiency nonnative speakers than oral interviews can offer.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Vocabulary profiles; TESL trainees; Academic success

* Corresponding author. Tel.: +1-514-978-3000x7649; fax: +1-514-978-4652.
E-mail addresses: morris.lori@uqam.ca (L. Morris), cobb.tom@uqam.ca (T. Cobb).
1. Introduction

University-based programs in teaching English as a second or foreign language (TESL/TEFL) are costly to run in terms of the demands they make on resources at a variety of different levels. Many teacher education courses operate optimally with low student-professor ratios, particularly in microteaching and supervised practica. They are also expensive in terms of the administrative time required to find and maintain internship placements. The selection of candidates who are likely to do well academically, require minimal administrative support in the course of their studies, and graduate on time is therefore extremely important. For this to happen, program administrators need to be aware of valid and reliable measures of TESL/TEFL student performance and have some means of testing for them in a cost-effective manner.

One measure that has been found to reflect linguistic and academic ability and be easy to administer is a computer-based analysis of vocabulary proficiency (Laufer, 1992; Laufer and Nation, 1995). The goal of the present study was to investigate the assessment potential offered by one such measure, Vocabprofiler, an on-line adaptation of Nation and Hwang’s vocabulary assessment instrument (Cobb, 2002). Our intention was to gain new insight into two different aspects of vocabulary profiling. Firstly, we wanted to measure the potential offered by Vocabprofiler for use in second language teaching trainee assessment. To this end we used entrance exam opinion essays to establish vocabulary profiles for 122 TESL trainees studying in a Canadian university and correlated the results obtained with the grades these students achieved in two academic courses of their program in a bid to measure the predictive power of the instrument. Our next goal was to evaluate the suitability of Vocabprofiler in assessing linguistically heterogeneous populations including both native speakers (NSs) and nonnative speakers (NNSs). Given the present-day make up of the ESL/EFL teaching community, it is safe to assume that any effective assessment instrument of English-teaching suitability should be able to evaluate both those who have English as a mother tongue and those who do not.

2. Research in the field

This study draws on findings from three different research strands: lexical knowledge, literacy, and metalinguistic awareness. These strands are woven across a variety of disciplines, including psychology, semiotics, applied linguistics and education, and involve studies that have included participants of different mother tongues, ages, nationalities and academic goals.

2.1. Lexical knowledge

The first strand to be considered is research into lexical knowledge and, more particularly, into current applications of vocabulary profiling. Work in this field is fairly recent in that it was born of technological advances of the past two decades.
that have made on-line analysis of large corpora feasible. As substantial amounts of English writing became available in electronic form and microprocessors gained power, researchers could conduct word frequency studies and establish a variety of different words lists. With some of these lists, it became possible to identify vocabulary needs by field, and this information proved to be extremely valuable in the teaching of English for academic purposes (EAP) and for special purposes (ESP). Furthermore, these lists provided a means of situating the word knowledge of any individual with respect to the depth and breadth of word knowledge needed to access various types and levels of writing. Such measures can be used to assess readiness for study or work in a second language. They can also serve to identify individuals in need of lexical remediation.

Traditional profiles, such as those pioneered by Nation, Laufer and a variety of different colleagues, provide breakdowns that include percentages from the one-thousand most common words (K1), the next thousand most common words (K2), the Academic Word List (AWL), or its predecessor, the University Word List (UWL), and off list words (Coxhead, 2000; Coxhead and Nation, 2001; Laufer, 1989, 1992; Laufer and Nation, 1995; Laufer and Paribakht, 1998; Nation, 2001; Nation and Waring, 1997).

For the purposes of the present study, one further calculation was added to the profiling repertoire: the percentage of function words\(^1\) (F), with a function word being defined as any word belonging to the closed set of words playing a grammatical role (e.g. articles, prepositions, auxiliaries, relatives, pronouns, and so forth). Function words form a subset of K1 as all function words in active use in English are frequently occurring lexical items. In the case of NS production written production, function words account for somewhere between 45% (academic writing) to roughly 50% (informal writing) of the total number of words used. The reason for their inclusion as a separate result in this study was a finding from recent research by Morris and Tremblay (2002) indicating that the percentage of function words in texts produced by ESL learners was a good indicator of the proficiency level. The learners in question ranged from the high beginner level to the high intermediate level. Interestingly, the higher their proficiency level, the greater their reliance on function words, with the most proficient students showing function word percentages ranging from 53 to 60%. Given these results, we were interested in determining if function words had anything to teach us about the expressive vocabularies of our NS and NNS study population.

As far as applications are concerned, vocabulary profiles have been used to examine differences between native and non-native, younger and older, and more and less educated speakers. This work has resulted in the charting of L1 and L2 vocabulary acquisition over time (Schmitt, 2000). It has also served to establish target vocabulary lists and proficiency levels for ESP and EAP programs (Cobb and Horst, 2001; Coxhead and Nation, 2001). From these applications and the research that has supported them have come a number of findings relevant to the present

---

1 This list differs from Meara’s Level 0 word list in that it includes all of the grammatical words of the language and not simply the most common.
study. Arguably the most important of these is the key role played by knowledge of words from the AWL or the UWL to academic success for L2 students at the tertiary level (Coxhead and Nation, 2001; Cobb and Horst, 2001; Laufer and Nation, 1995; Laufer and Paribakht, 1998). Researchers have found a consistent pattern of good academic performance in the case of NNS students who are able to make good expressive use of academic words. These results led us to look for similar correlations in our own study.

2.2. Literacy

The second major strand underlying this study involves research conducted in the field of literacy by Corson (1985, 1997), Olson (1977, 1994) and Olson and Astington (1990), Astington (2000), Astington et al. (1988), and Astington and Jenkins (1999). Corson’s work investigates the importance of the mastery of medium to low frequency words of Greek or Latin origin to the academic success of many learners, including second language children in the Canadian primary and secondary school systems. He finds that they often fail to gain the same access as their native speaking peers to higher-level, academic words as they advance through the school system because they do not start from a lexical base which is broad enough to support efficient lexical acquisition of less frequent items. In this respect, Corson’s findings are not unlike those of researchers in the field of lexical acquisition who find mastery of the words of the Academic Word List to be crucial to success in higher education. In both cases, academic success is amenable to an ability to access and use a higher lexical register when the situation so dictates. Those students who are unable to change lexical gears and shift from an informal, conversation register to an analytical, academic register encounter more and more difficulties at school as the reading and writing demands increase.

The work of Olson and Astington develops along a line similar to that pursued by Corson, but adopts a somewhat different perspective on the problem of the role of vocabulary knowledge in literacy. It looks at the acquisition of higher-level language skills in terms of the evolution of the thinking skills of the individual. Olson proposes that as a child or a second language learner gains control of a greater variety of lexical items “there is a transition from utterance to text both culturally and developmentally and that this transition can be described as one of increasing explicitness, with language increasingly able to stand as an unambiguous or autonomous representation of meaning” (1994: 258).

The importance of Olson’s contribution is the link it makes between language skills and the types of higher level thinking skills that university students and future teachers need to be successful. The notion that linguistic knowledge provides scaffolding for higher order intellectual progress and metacognitive development is not merely intuitively appealing; there is evidence from mother tongue studies of normal and language impaired children that language and developmental progress share a bootstrapping relationship in which lexis, grammar and cognitive ability develop together (Bates and Goodman, 1997; Bates, 1999). The extent to which this relationship is maintained in adulthood is unknown but certainly worthy of exploration,
particularly in light of the fact that many items from the Academic Word List have a metalinguistic or metacognitive function in discourse.

2.3. Metalinguistic awareness

The final research strand to be considered picks up the metalinguistic thread and links it to teacher education. There are two focal points for research into the awareness of second language teachers. The first is found in Hong Kong and is best represented by the work of Andrews (1997, 1999, 2001), who examines the metalinguistic awareness of pre- and in-service ESL/EFL teachers in terms of both declarative and procedural knowledge. Andrews’ findings touch on two topics of relevance to the present study. The fact that his in-service participants fare better than their preservice counterparts when it comes to an ability to state grammatical rules suggests that time on task is important to the assimilation of metalinguistic knowledge and not just linguistic knowledge. A second finding—one which will be reexamined in light of the present study—was that nonnative speakers outperformed native speakers in grammatical explanation tasks. It must be noted, however, that the NSs had little experience in metalinguistic analysis before doing the task while NNSs had better background knowledge. The sampling size was also very small.

The second focal point for teacher language awareness research is found in Canada and is represented by the work of Morris (1999, 2002, in press) who has also looked at the grammatical knowledge of pre-service teachers, but with a view to identifying traits that correlate significantly with an ability to perform a grammatical explanation task. Results obtained to date have identified age and, to a lesser extent, mother tongue as predictors of performance on such a task. Unlike Andrews’ findings, the results of Morris’ work indicate that NSs enjoy an advantage over NNS when it comes to an ability to identify and explain grammatical errors. It is noteworthy that this finding came from a study in which all of the participants received exactly the same grammatical instruction.

The work on teacher metalinguistic awareness is significant in that it identifies two factors that are likely to influence the academic performance of pre-service teachers: age and degree of knowledge, both implicit and explicit, of the target language. While Andrews’ and Morris’ finding do not concur regarding the importance of being a native speaker to achieving metalinguistic understanding, both studies find that older, more experienced language teaching trainees and language educators do better on the metalinguistic front than their younger counterparts.

Without too much suspension of disbelief it is possible to identify some common themes in the various research threads examined thus far. One recurrent theme is the importance of academic words, many of which play metalinguistic or metacognitive roles in discourse, to success in academic endeavours. Another common finding is the advantage of time on task to all sorts of learning, be it linguistic or pedagogic in nature. The more exposure people have to words, the more they acquire. The more time they spend giving grammatical explanations, the more proficient they become. It would thus seem that both duration of learning and context of learning are factors worthy of consideration in any study of academic performance. With this in mind,
we undertook our assessment of the predictive potential offered by vocabulary profiling in the assessment of a heterogeneous population of TESL trainees.

3. Research design

3.1. Participants

The study involved 122 Canadian TESL trainees registered in either a 4-year BEd or a 1-year Certificate program in a Quebec university. The participants were admitted into their program of study on the basis of their academic record, a 1-h opinion essay exam and an oral interview. While both the BEd and Certificate programs were officially given at the undergraduate level, many of those enrolled had previous degrees in a variety of different disciplines and were therefore not true undergraduates. Of the 122 participants in the study, only 47, or 39%, did not have at least an undergraduate degree. All were bilingual or multilingual and all had native-speaker proficiency or near native-speaker proficiency in spoken English, and a level of proficiency in written English that ranged from good to highly skilled. In many respects, the people under study constituted an elite group for an undergraduate program.

Tables 1 and 2 provide a summary of basic participant demographics. The L1 and age variation are both typical of the Canadian university context. Such heterogeneity represents a pedagogic challenge in the classroom, but is an asset in research terms. The real world of ESL is not one in which monolingual English speakers teach uniform classes of learners. Assessment instruments need to be validated on diverse populations.

Table 1
Participant profiles: L1

<table>
<thead>
<tr>
<th>L1</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>53</td>
<td>43.4</td>
</tr>
<tr>
<td>English-French</td>
<td>4</td>
<td>3.3</td>
</tr>
<tr>
<td>French</td>
<td>28</td>
<td>23.0</td>
</tr>
<tr>
<td>Greek</td>
<td>13</td>
<td>10.7</td>
</tr>
<tr>
<td>Italian</td>
<td>6</td>
<td>4.9</td>
</tr>
<tr>
<td>Russian</td>
<td>5</td>
<td>4.1</td>
</tr>
<tr>
<td>Armenian</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>Spanish</td>
<td>2</td>
<td>1.6</td>
</tr>
<tr>
<td>Romanian</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Portuguese</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Amharic</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Hebrew</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Arabic</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Slovak</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Dutch</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Hungarian</td>
<td>1</td>
<td>0.8</td>
</tr>
</tbody>
</table>
### 3.2. Data collection

Vocabulary profiles were established for each participant in the following manner. Entrance exam essays were used as a source of written production. In order to preserve the validity and reliability of the essay element of the admissions test, the topics given to participants varied over the period under study. In all, six different topics were used. While the use of a single topic would have heightened the reliability of the present research study, it would have had an adverse effect on the reliability of the university’s assessment of TESL candidates. All candidates were given a statement with which they were to agree or disagree in an opinion essay. They were instructed to produce at least 500 words in the hour they were given to write. No aids were allowed. All of the statements required the candidates to think carefully in order to construct a coherent argument. For example, one group had the following prompt: Violence is sometimes justified in the achievement of peace.

All essays that reached at least 300 words in length\(^2\) were collected and keyed in to a word processing program. Proper nouns were recategorized as frequently occurring words and spelling mistakes were corrected so that the proportion of low frequency words would not be skewed upwards. Once the cleaning up operation was complete, the first 300 words of each text were entered into Vocabprofiler (Cobb, 2002). For each text the profiler calculated the type/token ratio (TTR) and the percentage of words of the text falling into the first thousand most common (K1), the second thousand (K2), the Academic Word List (AWL), off-list words (OL), and, function words (F).

The measure of academic success was obtained by using the marks achieved in two obligatory grammar courses taken by all of the participants. The first course, (G1), was an introduction to English grammar. This course focused on knowledge of a declarative type (*knowing that*), and was heavily descriptive (rather than prescriptive) in orientation. The second course, (G2), was a pedagogical grammar course that dealt with how to teach grammar to ESL or EFL learners. This course

---

\(^2\) This resulted in the elimination of 11 of the original set of 133 texts.
involved more reliance on both procedural knowledge (*knowing how*) and analytical skills than the first course.

Results from the two grammar courses were preferred to using Grade Point Averages (GPAs) for two reasons: they were common to all of the participants and had been obtained according to common evaluation criteria. In contrast, the participants’ GPAs were based on results in a variety of courses, not all of them in the field of TESL, not all evaluated in the same manner, and not all showing the same discriminatory power. For example, administrative problems in some new courses in the program of study resulted in all students being given either an A or a B. This had the effect of inflating the GPA of certain students.

The final step in the methodological process was to measure the correlations between various frequency ranges of the vocabulary profiles and the different grammar course results. Correlations were calculated using Pearson $r$, and the differences between NS and NNS performance were assessed with ANOVAs.

### 4. Findings

The first set of findings are results from the correlation analyses. Table 3 provides a summary of the significant correlations found to exist between elements of the TESL students’ vocabulary profiles and results in the grammar course of their program of study. The highest correlation, $r=0.37$, found was between words on the Academic Word List (AWL) and grades in the pedagogical grammar course (G2). Two other significant correlations were found, one between K1, the thousand most frequent words, and G2, and the other between F, function words, and G2. In both cases $r=0.34$.

It is clear that none of these correlations is sufficiently high to support an argument in favour of vocabulary profiles as a stand-alone instrument of TESL candidate assessment. Indeed, the highest correlation, while significant, accounts for a shade over 6% of total variance. However, these correlations are nonetheless noteworthy for two reasons: the pattern of statistical significance in the profile categories and the fact that all of the significant correlations occurred with grades in the course devoted to the teaching of grammar. Both of these points will be picked up in the discussion section.

As for differences between the performance of NSs and NNSs, Table 4 provides a statistical summary of the results. In every comparison documented in the table NSs held the upper hand. They got better grades and, when writing, relied less on the first thousand words and on function words, and made more use of words from the

<table>
<thead>
<tr>
<th>Factors</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWL/G2</td>
<td>0.37**</td>
</tr>
<tr>
<td>K1/G2</td>
<td>$-0.34^{**}$</td>
</tr>
<tr>
<td>F/G2</td>
<td>$-0.34^{**}$</td>
</tr>
</tbody>
</table>

** $P<0.01$, 2-tailed.
Academic Word List. In short, NSs produced opinion essays that were better prototypical representations of academic writing as defined by results achieved through corpus-based analyses.

An alternative representation of the comparative results of Table 4 is offered in Table 5. Here the goal was to show the percentage of NS and NNS vocabulary profiles that met a fairly modest academic writing standard. This standard was established by correlating all of the TESL student vocabulary profile scores with academic grades and determining the levels which resulted systematically in academic success (i.e. the required C average in a course of study). The levels arrived at were the following: a K1 score of less than 85%, an AWL score of over 5%, and an F score of under 50%. The participants were then sorted according to these norms. Mann–Whitney Tests confirmed the statistical significance of all of the differences between NSs and NNSs ($P < 0.05$).

The gap between NSs and NNSs is made very salient by this form of analysis. The NNSs were only half as likely as NSs to have the ideal results across the board. Quite clearly, the lexical playing field was not a level one for NS and NNSs, even though it would have appeared to have been had the analysis been limited to interviewer assessments of oral proficiency. Indeed, some of the interviewers were very surprised by the vocabulary profiling results, particularly in the case of those they had felt to be complete bilinguals with native-like English skills.

5. Discussion

5.1. Viability of Vocabprofiler as an assessment instrument

Returning to the research questions posed at the outset, we can now consider our findings in light of the goals of the study. As was previously mentioned, the
correlations found between vocabulary profile results and grades were low enough to preclude the possibility of using Vocabprofiler as a stand-alone assessment instrument of potential TESL candidates. However, the inclusion of vocabulary profiles as part of an assessment battery is worthy of consideration, particularly since there is evidence that they can provide information about candidates that is otherwise difficult to extract but important to obtain. For instance, many of the researchers discussed in the literature review agree that those learners who are capable of using AWL words or other markers of formal register, such as words of Greco-Roman origin, are generally more successful in school. It could therefore be argued that identifying students who can access a formal register in appropriate circumstances would be useful. Interviews do not allow for assessment of this type of knowledge as AWL words are not normally ones used in basic oral exchanges. It takes a formal academic writing task to prompt people to produce the type of words that will allow for an adequate assessment of their ability to access a higher lexical register.

It must also be borne in mind that for the purposes of this study vocabulary profiling was only done on the entrance essays of students who were accepted into a TESL program and went on to register in and complete courses. In other words, our study considered how well vocabulary profiling could measure the degrees of academic success of successful applicants. Had profiling been extended to the texts written by unsuccessful applicants, it is quite likely that its value as an assessment instrument would have been much greater. Indeed, a quick check of a sampling of texts from students refused admission showed K1 percentages of well over 90% and low AWL percentages. Additional evidence in support of this argument comes from vocabulary profiles obtained from a group of 34 NNSs enrolled in a similar program of study with lower linguistic and academic entrance standards. These students were found to have vocabulary profiles that differed substantially from those of the participants in the present study (Morris, in press). For example, none of them had a K1 result of less than 85% and only 15% of the group managed less than 90% reliance on K1 words. Similarly, only 18% of the group had an AWL result that topped 5%.

Vocabulary profiling also offers certain advantages over straightforward passive vocabulary tests. While the latter are easy to administer and score, their multiple choice format that is characteristic of many opens the door for some candidates to commit to memory the words that will result in a good score without acquiring the cognitive skills that underlie these same words acquired in a more natural way. Secondly, passive vocabulary measures do not allow access to the additional information that can be derived from the examination of words produced in the context of a piece of writing that has a communicative purpose and a target audience. For instance, the texts produced for profiling analysis can be considered in terms of type-token ratios, lexical collocations, and even something as basic as spelling.

There is also an argument to be made in favour of vocabulary profiling on the basis of the fact that the significant correlations found between profile scores and academic grades all occurred with the more procedurally oriented of the two courses considered. This suggests that vocabulary profiles are indicative of something other than straightforward declarative knowledge of grammar, a construct which is rela-
tively easy to measure (Andrews, 2001; Morris, 2002). Although it is impossible to determine precisely what correlates with what, our findings suggest that vocabulary profiles offer a means of assessing ability to apply grammatical knowledge and not simply possess it. The course that dealt with the teaching of grammar had a very strong procedural component and was taught as a task-based course in which the grades achieved reflected an ability to identify grammatical errors, classify them according to severity, and respond to them in student production.

5.2. The suitability of Vocabprofiler in the assessment of a heterogeneous population

The argument for using vocabulary profiling in the assessment of a heterogeneous population is a compelling one in light of the findings reported above. Vocabprofiler proved to be efficient in sorting high proficiency NNSs into different levels of language ability. It also detected different levels within the NS population. Few if any instruments that are comparably user-friendly and inexpensive can claim to offer as fine a sorting potential of linguistically proficient students.

Since vocabulary profiling can be used to identify different bands of proficiency within the high proficiency range, it can also be used to identify students who may be at risk of encountering academic difficulties in their program of study before the problems actually arise in the classroom. This might serve to lower dropout rates if remediation could be provided for the at-risk population. For example, if better knowledge and use of AWL words generally results in better academic performance, students showing weak productive ability in these words might be taught how to use them more effectively in one of the compulsory writing courses students are usually required to take.

Identification of at-risk students by means of vocabulary profiling might have helped some of the 14 participants in this study who dropped out before completing their program. Of these 14, 11 had K1 scores of over 88% and an AWL percentage of under 5%, and thereby had profiles that were below ideal standards.

6. Conclusions

The evidence presented in this study would indicate that vocabulary profiles offer good potential as predictors of the academic performance of TESL students if used in conjunction with more traditional forms of entrance assessment, such as interviews, academic records, and tests of grammatical knowledge. They are simple to run, cost effective, and able to get at information that interviews and measures of declarative language knowledge do not reveal. Furthermore, they would seem to be an excellent means of measuring the English language skills of high proficiency NNSs, a very difficult population to assess using traditional means.

In addition to providing evidence concerning the assessment potential offered by vocabulary profiling, the findings of this study lead us back to the reflections of Olson on the relationship between language skills and the coherent representation of thought. While the levels of correlation between expressive vocabulary and
academic performance found in the course of this study were not high, they were nonetheless statistically significant and might have been even more significant if extended to all applicants to the TESL program, included those who were refused. More importantly, they would seem to confirm the intuitively satisfying notion that a better knowledge of academic words and a better ability to access a formal academic register results in better performance in metacognitively demanding courses. To do well in the course devoted to the teaching of grammar, trainees had to be able to evaluate their own grammatical knowledge and that of ESL learners, and then make a series of decisions about what could and should be done in a classroom setting. In other words they had to develop a representation of grammatical knowledge and then present selected aspects of it to a target audience. The trainees who accomplished these tasks the best were usually the ones who showed the best command of AWL words and the richest expressive vocabularies. This gives us cause to reflect in more depth on the complexity of the relationship between lexical knowledge and thinking, and the consequences for university classrooms in general and those that train teachers in particular. We clearly need to know far more than we do, not just about the assessment potential offered by vocabulary profiling, but about the role of word knowledge in the representation and transmission ideas.

References


Morris, L. Linguistic knowledge, metalinguistic knowledge and academic success in a language teacher education programme. Language Awareness, 12, 2 (in press).


