

Classrooms as Lexical Environments

PM Meara *Swansea University*, PM Lightbown and RH Halter *Concordia University, Montreal*

Abstract

In many second/foreign language classrooms, students are expected to learn much or even most of their vocabulary without explicit instruction, simply through exposure to a rich variety of words in meaningful contexts. In fact, however, there are few studies which would allow us to estimate the number of words learners are typically exposed to in second/foreign language classrooms. In this study, the vocabulary available in the speech of ten teachers in intensive communicative ESL classes for children in Quebec was analyzed using specially designed computer programs. The words which occurred in classroom transcripts were classified according to their status as high frequency or 'unusual' words, according to lists developed by Nation (1986). The working assumption was that a large number of unusual words would be indicative of a rich lexical environment whereas the absence or extreme rarity of such words would indicate that the classroom vocabulary was poor. The number of unusual words was found to be quite low in short periods of classroom interaction. However, an interpretation of the findings suggest that the actual richness of the vocabulary available may be greater than it appears in terms of this measure.

Introduction

This paper is an exploration of the vocabulary available in some English second language (ESL) classes in which teachers have made a strong commitment to a communicative approach to language teaching. It is also an account of some problems that we ran into when we attempted to establish just how rich a lexical environment language classrooms can provide.

After a long period of focussing mainly on syntax and morphology, second language acquisition (SLA) researchers have begun to pay more attention to the development of vocabulary (see, e.g., Meara, 1984; Gass, 1988; Harley, 1995). Language teaching practices have changed considerably in recent years, and it is important to try to determine how vocabulary development takes place in language classrooms which reflect these changes. Many teachers who are committed to 'communicative language teaching' are reluctant to

Classrooms as lexical environments

provide extensive explicit teaching of words and their definitions or derivations or to assign students vocabulary lists to memorize. They try instead to encourage learning in context by providing learners with interesting classroom activities. In these activities, students are expected to hear and learn many new words, principally by hearing them within a context which permits their meaning to be inferred.

Although we suspect that there is still a valuable role for some explicit vocabulary building activities, we believe it is reasonable to assume that a great deal of vocabulary, like syntax and morphology, can be acquired through comprehensible input especially when the oral input is augmented by reading (see e.g., Krashen, 1989, 1990, 1993). An unanswered question, however, is just how much available vocabulary there is for learners whose exposure to their second language takes place principally in foreign language classrooms.

In this study we examined transcripts of language classrooms and carried out some semi-automated word-counts on these transcripts. Such word-counts can give us some idea of the richness of the vocabulary being used by the teachers of these classes. A large number of different words, spanning a wide frequency range, would suggest that the teachers were providing a relatively rich lexical environment for their students, while a small number of different words, drawn mainly from the high frequency words, might suggest that the teachers were simplifying the language they used to match the lexical abilities of their students. This might result in slow vocabulary growth, as learners hear the same high frequency words over and over again. Data from transcripts would allow us to answer questions like the following:

How many new words does a learner typically encounter in a class period?

How often are these new words repeated over a given time span?

Does lexical richness depend on the type of class being taught?

Is lexical richness affected by the level of the students? by their L1 background?

Questions of this sort are not trivial, but we have been able to locate very few studies which are concerned with the lexical environments that foreign language learners in modern foreign language classrooms operate in. We have, however, found some suggestive studies which address the question indirectly.

Scholfield (1991), examined a series of English language text-books from a lexical point of view. He was particularly concerned with the rate at which new vocabulary is introduced by course writers. His analysis demonstrates that some writers are much more systematic

Classrooms as lexical environments

about the way they introduce new vocabulary than other writers are. Some authors, for instance, introduce a small number of items in each new lesson, and systematically recycle old words so that they are not forgotten. Other authors appear to adopt a completely haphazard approach to vocabulary building, sometimes introducing huge numbers of new items in a section, sometimes a much smaller number, apparently without any consideration of what an optimum input rate might be. Scholfield's analysis does not include a consideration of what teachers do to make the lexical load of their textbooks more manageable, but the implication of his work is that some learners might be operating in lexical environments which are, if anything, too rich for them to handle.

Another relevant piece of work is an informal account of an analysis of several sets of radio programs aimed at learners of English at a variety of proficiency levels (Meara 1993). The programs were broadcast by BBC English, the English teaching section of the BBC World Service. The programs examined in this study included some which were intended for beginners as well as some for learners at higher levels of proficiency. Meara found that radio broadcasts aimed at learners of English had a surprisingly small range of vocabulary. He reckoned that these broadcasts included one 'unusual' word every two or three minutes of broadcast time, where 'unusual' is defined as words which are not among the 2500 most frequent words in English. Meara also showed that over a whole series of related programmes, unusual words tended to get repeated, so that the number of genuinely new and unusual words became steadily smaller. These words were often explicitly taught, so that the real lexical load of these programmes was reduced even further. Meara concluded that it was difficult to see how these programmes might make a significant contribution to the lexical repertoire of a listener at intermediate level.

Henzl (1973) carried out a study which was in a sense a simulation of the classroom environment. In this study, she asked eight native speakers of Czech to tell two stories to (1) American university students who were studying Czech and (2) a group of native speakers of Czech. Transcriptions were made of speech samples recorded when the native speakers told a political anecdote and a descriptive story to each of these audiences. Henzl characterized the linguistic register of the speech addressed to the two groups. She called the register which typified the stories addressed to the university students a foreign language classroom register (FLCR), while the versions addressed to the native speakers were in colloquial Czech (CC). Henzl's analysis of the lexicon in these samples led her to the following observation:

One of the main characteristics of the FLCR samples seems to be the great use of basic vocabulary, while native discourse, as a rule, exhibits an extremely rich diversity in the choices of words... The variety of speech in most FLCR samples lacked many words and phrases found frequently in CC, and [CC], for various reasons, stayed at the periphery of the SLC [standard literary Czech] lexicon. (p. 210)

Classrooms as lexical environments

Second- and foreign-language teaching has changed considerably in the years since Henzl's study. In some ways, however, her research using a story-telling technique (the presentation of comprehensible input) is more relevant to current teaching practices than some of the studies which might have been done in actual classrooms of the period. In the 1960s, language teaching in North America was dominated by audio-lingual methods. Theorists supporting this approach recommended the limitation of vocabulary in order to allow learners to concentrate on manipulating 'the structures' of the language (see e.g., Brooks, 1964). Current 'communicative' and comprehension-based approaches to second and foreign language teaching, in contrast, encourage the use of authentic materials such as television and radio broadcasts and relatively free interaction among learners engaged in a variety of tasks and projects (see, e.g., Brumfit and Johnson, 1979; Long and Crookes, 1993). It seems likely that such communicative approaches would lead to the availability in the classroom of a far richer vocabulary than was the case in the classrooms which were typical of approaches which emphasized 'structure' over vocabulary.

Suggestive though they are, neither of these pieces of work throws any light on classrooms as lexical environments. In a search of the literature, we have not yet found empirical studies which specifically investigated the lexical resources that are available to students in foreign language classes. There is no doubt that the relationship between teaching methods and materials and the vocabulary actually available in the classroom will be complex. Since current teaching practices are based on the belief that learners will acquire much of their vocabulary from the comprehensible input available in the classroom, it is clearly important to know what vocabulary is likely to be available there. Many teaching programs, for example, no longer use textbooks, preferring to base their lessons on more spontaneous language which arises as needed in the range of activities which teacher and students engage in. Thus, the only way to know what vocabulary students are likely to have access to is to analyze the language which is actually used in classrooms. There is no satisfactory substitute for actual classroom data.

1. Approach

In this study, we set out to answer the questions listed earlier. We initially planned to carry out our analyses on existing transcripts from a variety of language classrooms. This would permit us to do the simple word counts which in turn would provide some preliminary answers to our research questions. Recent years have seen numerous publications on classroom language acquisition (e.g., Ellis 1990; Allwright 1988; Allwright and Bailey 1991), and we assumed that the analyses reported in these books would be based on transcripts. This assumption turned out to be incorrect. Much of the data quoted in classroom interaction research is based on observational studies using real time coding schemes or partial transcriptions from specific classroom events. One exception to this is Chaudron (1982). Chaudron used transcripts of a number of ESL classes to study the ways in which teachers were explicitly trying to help learners understand vocabulary in the

Classrooms as lexical environments

classroom. Håkansson (1986) also analyzed classroom transcripts and determined, for example, how many word (tokens) per minute were uttered by teachers. However, none of the studies we found had asked the questions about available vocabulary in the classroom which we wanted to explore, and there was no large corpus of classroom transcripts available for our analysis. A query posted on the Second Language Acquisition and Research bulletin board, SLART-L, requesting access to classroom transcripts, yielded no responses. Recently, we have managed to locate a number of small transcript corpora, and some of this material has been made available to us by generous colleagues. For this paper, however, we decided to analyze a sample of transcripts from some of our own previous research. The transcripts analyzed in this study were made in the context of an innovative ESL program in Quebec. Students participate in this program when they are 11 to 12 years old. Their first language is French, and they live in neighborhoods or regions where they rarely come into contact with any language other than French. In grades 4 and/or 5, they have ESL classes one or two times a week (90 to 120 minutes per week in all). In grade 5 or 6, they participate in "intensive" ESL classes. For five months of this one school year, they spend most of every school day in a variety of activities designed to help them learn English (Spada and Lightbown 1989; Lightbown and Spada 1994). The classes in this program are characterized by a variety of "communicative" activities. In most of the many classes we have observed, there is virtually no explicit teaching of grammar or vocabulary and very little error correction. Students do many tasks and projects in groups while the teacher offers assistance as needed. Roughly half of the classroom tasks are teacher-centred.

The corpus analyzed for this paper consists of 10 samples of just over 30 minutes of class time each. Each sample was produced by a different teacher, and all were recorded near the end of the 5 month intensive program. We can estimate that students had had about 400-500 hours of English instruction -- about 350 of them in the intensive program. The complete recordings covered many more hours of classroom activity than those analyzed here. In choosing portions to be transcribed, we consulted information obtained in real-time coding using the Communicative Orientation of Language Teaching (COLT) classroom observation scheme (Spada and Fröhlich, 1995). From the COLT coding sheets, we identified periods of classroom activity which were relatively teacher-centered. There were two reasons for this. The practical reason is that transcription of learner-centered activities, especially those which are carried out by students in groups, is extremely difficult, and our recordings did not permit us to reliably transcribe the learners' language in such activities. In addition, it was assumed that the teacher is the principal source of new or unusual words, the focus of our study here. It should also be noted that none of the transcripts analyzed contains activities in which the teacher explicitly focussed on teaching vocabulary.

For this paper, then, we return to the first of the questions listed above: How many new

Classrooms as lexical environments

words does a learner typically encounter in a class period?

This question is not as straightforward as it looks at first sight. It is obvious, for instance, that one cannot just run an unedited transcript through a naive word-counting program in order to answer questions of this sort. After a number of false starts, we made the following operational decisions:

1: For the basic analysis we discarded the student data and examined only the input from the teachers.

2: Removing the learner data left us with transcripts that varied considerably in length. Since all the major measures of lexical richness are affected by the length of the sample, we standardized our samples at 500 word tokens. We were able to extract at least two samples of this length from all but one of our transcripts.

3: For the purposes of this analysis, we started from an assumption that the students had a vocabulary of about 2500 basic words. This figure was taken from Nation (1986), which is a list of words that foreign language learners can be expected to meet early in their learning experience. Nation's lists are largely based on earlier frequency counts, notably West (1953) and Kucera and Francis (1967), but they also take into account factors affecting the vocabulary that learners are exposed to. For example, the lists include items which are common in many language classrooms ("noun", "verb", "sentence", etc.) and items which learners commonly acquire as closed sets (days of the week, and so on). For practical purposes, Nation's list can be subdivided into 4 levels: NAT0 comprises about 500 highly frequent closed class words such as articles, prepositions, common greetings and courtesies, numerals and so on. NAT1 words correspond to the one thousand most frequent words in English (apart from those in NAT0), supplemented by common objects in the local environment. NAT2 words correspond to the second thousand most frequent words in English. We assumed that learners in this study would know most of the words in NAT0, NAT1 AND NAT2. The list also includes a set of words (designated as NAT3) which is largely concerned with common scientific and technological concepts. The NAT3 list is designed to include words which students about to begin post-secondary education in English would be expected to know.

4: Our working assumption was that an environment which was restricted to the basic vocabulary could be characterized as 'lexically poor'. A lexically rich environment would be one which included a larger proportion of unusual words. It must be understood that 'poor' and 'rich' are relative terms. For absolute beginners, any target language environment would be challenging. Obviously, a lexical environment which was made up of too many 'unusual' words would make it difficult or impossible for learners to infer the meaning of a new word from context. However, we do not know what the optimal proportion of known to unknown words would be.

Classrooms as lexical environments

5: It is conventional in studies of lexical richness to distinguish between *word types* and *word tokens*. We elected to base our analyses on *word types*. Type counts ignore the number of times a word is used in a text, and simply identify the number of different words. Unusual words will generally appear only once or twice, whereas more ordinary words tend to appear several times, and this has some important consequences for the way we evaluate lexical richness. Generally speaking, using type counts makes the number of unusual words appear more important. For example, if we take a text of 100 tokens, and find five unusual word tokens, we have 5% 'unusual words'. However, when account is taken of repetitions, then the same text might only consist of 50 different word types, and the proportion of unusual types in the text is then five types out of 50, or 10%.

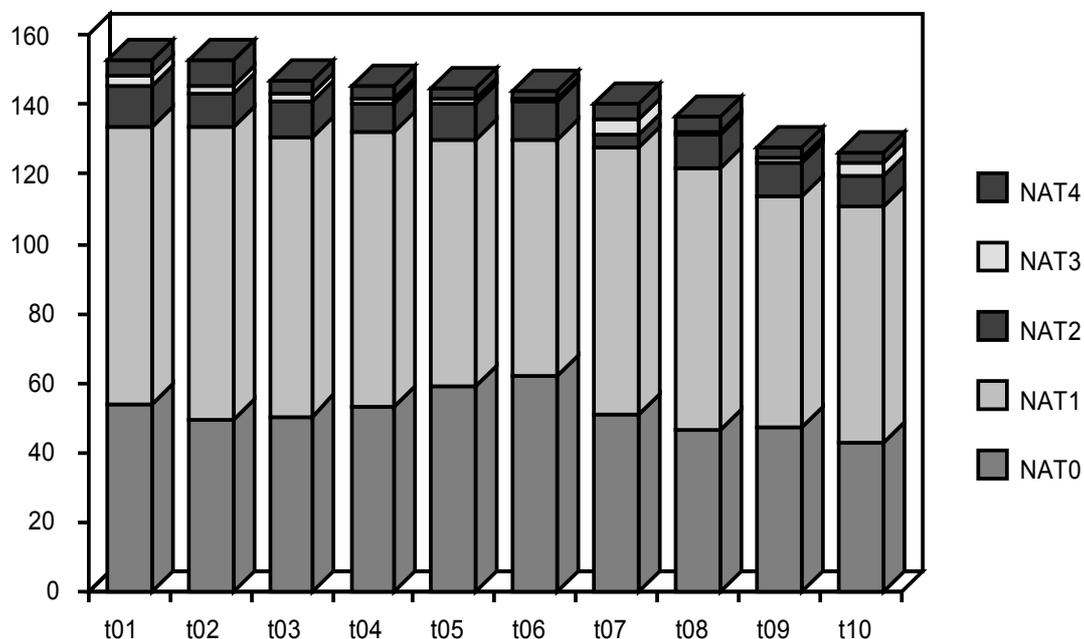
6: It was also necessary to decide how derived forms of simple words are going to be handled. The easiest way to analyze texts using computers is simply to count as a different word form any string of letters that is separated by spaces. This simple solution raises a number of obvious problems. For example, this would lead us to count TALK, TALKS, TALKED, TALKING as four different words. We decided to count not words but lemmas. In a lemma count, different forms of "the same" word are counted as instances of a single lexical item. That is, the forms of *talk* are all treated as instances of the base lemma TALK. Because many frequent words are derived forms, lemmatizing the word lists has the effect of reducing the number of ordinary words, and thus increasing the proportion of unusual words.

2: Procedure

The texts were processed using specially written computer programs. These programs take as input a raw transcript and produce from it an unedited word list and a lemmatized word list. The programs then identify which of these lemmas are found in which of Nation's lists, categorizing each lemma as NAT0, NAT1, NAT2 or NAT3. All proper names are categorized as instances of NAME, and all numbers as instances of NUMERAL. All other items which cannot be found in Nation's lists are considered to be 'unusual' and are assigned to a category labelled NAT4.

This analysis permits us to determine how many unusual words (defined as words in the NAT4 list) occur in the transcripts, and what proportion of the total lemma count they make up. The raw data are shown graphically in Figure 1 (See Appendix 1 for the data Figure 1 is based on). This figure shows the mean number of different lemmas used by nine teachers in two 500 word samples and by one teacher (t06) in a single 500 word sample. The figure also shows how these lemmas are categorized according to Nation's lists.

Figure 1. Lemma Types: Mean of Two 500 Word Samples for Each Teacher



The most striking thing about Figure 1 is how very homogeneous the data are. The overall mean number of lemmas is 141.65 and all of the 10 teachers are very close to this figure. Also striking is the similarity of the distribution of the lemmas in each data set. In all 10 cases, about one third of the lemmas come from NAT0 (36%), and around half of the lemmas come from NAT1 (53%). NAT2 words account for a further 7%. NAT3 appears to account for one or two lemmas in each sample. The remaining 3% are the lemmas that we have classified as unusual. Again, all the teachers lie very close to this average figure. At first glance, these data seem to suggest that these classrooms are not rich lexical environments. The number of unusual words in any single sample is very low, and it may be that even these very low figures might be overestimating the number of unusual words occurring in the texts. Many of the words identified as unusual have close cognates in French. Since all students in this study were native speakers of French, one might expect that such words ought to be easily recognized and easily learned by the students. Examples of this sort include FANTASTIC, MUSKETEER, ANNIVERSARY, SOUVENIR and BOULEVARD. If we exclude these words from our list of unusual items, then the number of unusual words in each sample is reduced by about half (See Table 1).

Once again, the homogeneity of the data is striking. Only one teacher, t02, has an average greater than 3.5 unusual words per sample. On average, it appears that genuinely unusual

Table 1: Unusual words in each 500 word sample (excluding cognates)

	t01	t02	t03	t04	t05	t06	t07	t08	t09	t10
smpl1	0	4	6	2	1	2	3	3	3	0
smpl2	5	7	0	3	1	-	2	4	2	2
mean	2.5	5.5	3.0	2.5	1.0		2.5	3.5	2.5	1.0

words appear in these transcripts at a rate of about 2.75 per 500 words of running text.

3 Discussion

Our initial reaction to these data was that the classrooms do not look like rich lexical environments. However, there are a number of factors which suggest that this interpretation might not be the correct one.

Each of the 500 word samples represents about 15 minutes of class time. This means that, at this rate, a normal class of fifty minutes would contain no more than 10 words drawn from outside the 2500 basic vocabulary we started with. If these classes were once-a-week classes, then it would indeed be legitimate to conclude that the lexical environment for these learners was very poor. Even over the course of a year, learners would come across only 400 unusual words from this source. However, the learners in the classes we analyzed are not attending once-a-week classes: in intensive classes, they are exposed to input of this sort for about five hours a day five days a week. If we assume (conservatively) that the observed rate of unusual words is typical of all the classes that these students attend, and that repetition does not reduce the input over time, then over the course of a single day they could be exposed to something like 50 unusual words - or 250 unusual words in a week. Suddenly, the lexical environment begins to look very much richer.

There are other reasons to believe that our analysis may systematically underestimate the lexical richness of these environments. The first is that we assumed throughout this analysis that the learners already knew the 2500 words from Nation's first three word lists. This assumption is not necessarily true, of course. There is a strong case to be made for adopting a much more conservative assumption that these learners have a good grasp of only a very small basic vocabulary.

Johnson and Swain (1994) cite studies showing that students of a similar age group whose ESL experience is somewhat comparable to that of students in this study also have a smaller than expected vocabulary. These are students in 'late immersion' courses in Hong Kong will have had limited amounts of ESL instruction in the preceding primary school

Classrooms as lexical environments

years.

In a study of textbooks used in the first year of late immersion (grade 7), Cheung (1985) showed that the textbooks contained the vocabulary of 2-3000 head words (5-6,000 words) and that these, and the sentence structures used were well beyond the students' reading ability. Later studies (Education Department, 1988; Lai, 1991) showed that approximately 15% were unable to read the simplest of simplified English readers and a few students were able to read and levels requiring a vocabulary of more than 500 words (Johnson and Swain, 1994, 218-219).

Before they entered the intensive ESL class, these students had been exposed to a total of less than 150 hours of English, spread out over grades 4 and 5. During the intensive class itself, they would have had approximately 350 more hours of exposure. There is no clear indicator of how many words should have been acquired by learners with a total of 500 hours of 'communicative' exposure to English. However, if these learners reliably know only the NAT0 and NAT1 level words, then for them the number of 'unusual' words occurring in the transcripts would more than double.

It is difficult to assess the force of this argument since we did not analyze the students' vocabulary knowledge. We know, however, that conservative estimates of vocabulary knowledge at this level of proficiency are more likely to be accurate than more generous ones. We were not able to assess the vocabulary of the learners who were in the classrooms studied here because the classroom recordings were made several years ago and over a period of several years. However, we were able to collect some vocabulary size data from a large group of learners in comparable learning situations and with comparable backgrounds in terms of native language (French) and prior exposure to English (Lightbown and Spada, 1996; White, 1996). In 13 classes, approximately 390 students were tested, using vocabulary tests covering the first 1500 words in English (Meara, 1992), i.e., levels NAT0 and NAT1. This group of learners, who were at a level very similar to the level of learners in the classes analyzed here, obtained a mean score of about 66% (s.d. 17.11). There is a considerable range around these mean scores, with some individual learners scoring near 90% and others scoring below 20%. These data lead us to conclude that many of the students in the classes must have had significant gaps in their knowledge of the 2500 basic words of English which we had originally assumed they might already know. Indeed, the data suggest that we can reliably assume that many learners at this level do not know even some of the basic words in the NAT0 and NAT1 lists. Without further detailed studies with individual subjects, it is difficult to assess how serious their vocabulary gaps might be. In any case, it suggests that for any individual learner, the number of unknown words appearing in the teachers' speech on a given day may be very substantial indeed.

Another reason to reconsider the actual richness of the available vocabulary in these

Classrooms as lexical environments

classrooms is the way we treated cognates in this analysis. We excluded from our list of unusual words close cognates which ought to be transparent to French speakers. However, it does not follow that all the items we identified as cognates would be readily recognized as such by the students. This is especially true in oral input where resemblances between words in different languages may be harder to perceive. In any case, even if cognates are easier to learn than non-cognate items, they still have to be learned, and this may mean that it was unwise to exclude them completely from our counts. It seems safe to assume that some of the cognate words in our transcripts are words that the students had not yet met as English words, even if they were familiar with these forms in French.

The third reason to reassess the richness of the vocabulary in these classrooms, concerns the way we treated derived forms. In our counts, we reduced all derived forms to their base lemmas, and classified these words according to the frequency of these roots. This procedure was justified on the grounds that most morphology in English is productive, and that students can learn to decompose the morphologically complex parts into their roots once they are familiar with the way English morphology works. On reflection, however, it is not obvious that this argument applies to all derived forms.

It probably makes sense to treat *LOVES*, *LOVED*, and *LOVING* as instances of *LOVE*, but forms like *GO*, *WENT*, *GONE* or *BE*, *AM*, *IS*, *ARE*, *WAS*, *WERE* seem more problematic. There is a case to be made for treating forms like these as separate lexical items, at least for beginner level learners. In the context of L1 acquisition, Prasada and Pinker (1993) have argued that, for children, regular verbs (e.g., *LOVE*) or nouns with regular plurals (e.g., *TREES*) may be stored as the base form and altered as needed by the application of a rule while irregular forms (e.g., *GO* and *WENT*) are actually stored as separate lexical items.

A related argument has been put forward in second language situations by Bauer and Nation (1993) (see also the discussion of this point in Laufer and Nation 1995). Bauer and Nation suggest that morphological patterns are gradually absorbed into learner's lexical repertoires. They see a strong case for counting morphologically complex words as separate items as long as the learners are at a low proficiency level. Specifically, for example, they have suggested that different lemmatization rules would be appropriate for different contexts. For example, *SING*, *SINGS*, *SANG*, *SINGER* would count as four words at level 1; *SING/SINGS*, *SANG*, *SINGER* would count as three words at level 2; *SING/SINGS/SANG/SINGER* would count as a single lemma at level 3, and so on. Our analysis has very largely ignored the possibility that learners might not perceive *AM* and *WERE*, for example as variants of *BE*, but rather as quite separate words. Inspection of the raw word lists suggests that although derived forms do not play a very large part in our data, their contribution is not altogether insignificant either.

Classrooms as lexical environments

Taken together, these three considerations suggest that the analyses we presented earlier might underestimate the true lexical richness of the classroom environments of these learners. The range of new words encountered in any single 15-minute period may not seem particularly high. Cumulatively, however, the number of new words likely to be met by students in intensive ESL classes may be more than we would expect them to be able to learn. Furthermore, as noted above, there is reason to believe that they may not already have mastered the 2500 words we originally thought they might know.

4. Conclusion

Clearly, the data reported here do not provide conclusive evidence that classrooms can be characterized as rich lexical environments. The transcripts we analyzed are not typical of most classroom language learning. This study was based on intensive L2 classes, providing several hours a day of exposure to the second language, a luxury which is not available to most second and foreign language learners. Of course, this does not mean that the students do in fact learn all or most of the words available to them in this environment. It is quite possible that despite the possibilities available to them, learners in these intensive L2 environments are still unable to acquire large vocabularies without greater access to written texts or to more explicit instruction. If this is true, however, the problem is not that the classrooms are too impoverished from a lexical point of view. Rather, the problem may be that students' ability to identify, understand, and learn the available vocabulary is limited. Krashen (e.g., 1989, 1992) suggests that the best way for second language learners to develop their vocabulary in the second language is through reading, especially free voluntary reading. Others have suggested that, within the oral interactions in the communicative classroom, teachers need to give students more guidance in learning and remembering words in the 'available' vocabulary (see, Chaudron, 1982; Clipperton, 1994; Lapkin and Carroll, 1987; Swain and Carroll, 1987).

We are currently carrying out a study of the lexical richness of other classrooms, contrasting the vocabulary made available by audio-lingual and communicative approaches to the teaching and learning of a second language (Lightbown, Halter, and Meara, 1999). In addition, future studies will examine the effects of treating derived forms of the same lemma as different word types. We also hope to return to the question of the students' learning of the words that are available to them.

5. References

Allwright, R.

Observation in the language classroom. London: Longman. 1988.

Allwright, R. and KM Bailey.

Focus on the language classroom. Cambridge: Cambridge University Press.

Bauer, L. and ISP Nation.

Word families. *International Journal of Lexicography*, 6(1993), 253-279.

Classrooms as lexical environments

Brooks, N. (1964).

Language and Language Learning: theory and practice, second edition. New York: Harcourt, Brace. 1964.

Brumfit, C., and K Johnson.

The Communicative Approach to Language Teaching. Oxford: Oxford University Press. 1979.

Chaudron, C.

Vocabulary elaboration in teachers' speech to L2 learners. *Studies in Second Language Acquisition*, 4(1982), 170-180.

Clipperton, R.

Explicit vocabulary instruction in French immersion. *Canadian Modern Language Review*, 50, 736-751.

Donoghue, M.

Foreign languages and the elementary school child. Dubuque, Iowa: Wm. C. Brown. 1968.

Ellis, R.

Instructed Second Language Acquisition. Oxford: Blackwell. @@@

Gass, S.

Second language vocabulary acquisition. *Annual Review of Applied Linguistics*, 9(1988), 92-106.

Håkanson, G.

Quantitative aspects of teacher talk. In **G. Kasper** (Ed.), *Learning, teaching and communicating in the foreign language classroom*. Aarhus: Aarhus University Press. 1986.

Harley, B. (Ed.).

Lexical issues in language learning. Ann Arbor, Mi.: Language Learning / John Benjamins. 1995.

Henzl, VM.

Linguistic register of foreign language instruction. *Language Learning* 23(1973), 207-222.

Krashen, SD.

Principles and practice in second language acquisition. Oxford: Pergamon. 1982.

Krashen, SD.

The input hypothesis: Issues and implications. London: Longman. 1985.

Krashen, SD.

We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. *Modern Language Journal* 73(1989), 440-464.

Krashen, SD. and TD Terrell.

The Natural Approach. Oxford: Pergamon Press. 1983.

Kucera H. and WN Francis.

A Computational Analysis of Present-day American English. Providence, RI.: Brown University Press. 1967.

Lado, R.

Language teaching: A scientific approach. New York: McGraw-Hill, Inc. 1964.

Classrooms as lexical environments

Lapkin, S. and S Carroll.

The role of vocabulary in the development of communicative ability: Some comments. Proceedings from the Symposium on the Evaluation of Foreign Language Proficiency. Bloomington: University of Indiana. 1987.

Laufer, B. and ISP Nation.

Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics* 16(1995), 307-322.

Lightbown, PM., R Halter and PM Meara.

Contrasting patterns in classroom lexical environments. In: **D Albrechtsen, B Henriksen, I Mees and E Poulsen** (Eds.) *Perspectives on Foreign and Second Language Pedagogy*. Odense: Odense University Press. 1999.

Lightbown, PM, and N Spada.

An innovative program for primary ESL in Quebec. *TESOL Quarterly*, 28,3(1994), 563-579.

Lightbown, PM, and N. Spada. Learning English as a second language in a special school in Quebec. Manuscript. 1996.

Long, MH. and G Crookes.

Units of analysis in syllabus design: The case for task. In **G. Crookes and SM. Gass** (Eds.), *Tasks in pedagogical context. Integrating theory and practice*. Clevedon, Avon: Multilingual Matters. 1993.

Meara, PM.

The study of lexis in interlanguage. In **A. Davies, C. Criper, and A. Howatt** (Eds.). *Interlanguage* (pp. 225-235). Edinburgh: Edinburgh University Press. 1984. 225-235.

Meara, PM.

EFL vocabulary tests. Centre for Applied Language Studies, University College Swansea, 1992.

Meara, PM.

Tintin and the World Service, Hornby Trust Lecture. IATEFL Annual Conference report, 1993. 32-37.

Nation, ISP.

Word lists: Words, affixes and stems. (revised edition) Wellington: Victoria University English Language Centre. 1984.

Prasada, S., and S Pinker.

Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes* 8,1(1993), 1-56.

Scholfield, P.

Vocabulary rate in course books: Living with an unstable lexical economy. Thessaloniki: Aristotle University. 1991.

Spada, N., and M Fröhlich. (in press).

The Communicative Orientation of Language Teaching (COLT) Observation Scheme: Applications for research and teacher education. Sydney: NCELTR.

Spada, N. and PM Lightbown. Intensive ESL programmes in Quebec primary schools.

Classrooms as lexical environments

TESL Canada Journal 7(1989), 11-32.

Swain, M. and S Carroll.

The immersion observation study. In **B. Harley, J.P.B. Allen; J. Cummins, and M. Swain** (Eds.). *The development of bilingual proficiency: Final report. Vol II: Classroom treatment.* Toronto: Modern Language Centre, Ontario Institute for Studies in Education. 1987.

West, M.

The general service list of English words. London: Longman, Green, and Co. 1953.

White, J.

An input enhancement study with ESL children: Effects on the development of personal pronouns and possessive determiners. Unpublished PhD dissertation, McGill University, Montreal. 1996.

A version of this paper appeared in *Language Teaching Research* 1,1(1997), 28-47.

Classrooms as lexical environments

Appendix 1

Distribution of Lemma Types: Mean of Two 500 Word Samples for Each Teacher (t06 = One 500 Word Sample)

	t01	t02	t03	t04	t05	t06	t07	t08	t09	t10
Nat4	4.00	7.00	3.50	3.50	3.00	2.00	4.00	4.50	3.50	3.00
Nat3	3.00	2.50	2.50	2.00	1.00	1.00	4.50	0.50	1.50	3.50
Nat2	12.00	9.50	10.50	7.50	10.50	11.00	5.00	10.00	9.50	8.50
Nat1	79.50	84.00	80.00	79.50	71.00	68.00	76.50	75.00	66.00	68.00
Nat0	54.00	49.50	50.50	53.00	59.00	62.00	51.00	46.50	47.50	43.00
Total	152.50	152.50	147.00	145.50	144.50	144.00	140.00	136.50	128.00	126.00

Classrooms as lexical environments