



PROJECT MUSE®

---

## From Faible to Strong: How Does Their Vocabulary Grow?

Marlise Horst, Laura Collins

The Canadian Modern Language Review / La revue canadienne des langues vivantes, Volume 63, Number 1, September/septembre 2006, pp. 83-106 (Article)

Published by University of Toronto Press

DOI: <https://doi.org/10.1353/cml.2006.0046>



➔ *For additional information about this article*

<https://muse.jhu.edu/article/204197>

# From *Faible* to Strong: How Does Their Vocabulary Grow?

---

Marlise Horst  
Laura Collins

**Abstract:** The study drew on an 80,000-word corpus consisting of narrative texts produced in response to picture prompts by 210 beginner-level francophone learners of English (11–12-year-olds). The unique feature of the corpus is its longitudinal character: The samples were collected at four 100-hour intervals of intensive language instruction, during which time students made considerable progress in listening and speaking. However, analysis of these staged sub-corpora using Laufer and Nation's 1995 Lexical Frequency Profile did not identify the expected increase in use of less frequent words. Further analyses using three measures available at [www.lexutor.ca](http://www.lexutor.ca) (a Greco-Latin cognate index, a count of word families, and a types-per-family ratio) showed that although the learners continued to use large proportions of frequent words, their productive vocabulary featured fewer French cognates, a greater variety of frequent words, and more morphologically developed forms. Implications for frequency-based vocabulary acquisition research and vocabulary teaching are discussed.

**Résumé :** L'étude a été effectuée à partir d'un corpus de 80 000 mots tirés de textes narratifs produits en réponse à des images incitatives par 210 francophones apprenant l'anglais (niveau débutant; 11-12 ans). Le corpus se distingue par son caractère longitudinal : les échantillons ont été recueillis à quatre intervalles de 100 heures de cours de langue intensifs, intervalles durant lesquels les étudiants ont fait des progrès considérables pour ce qui est de leurs aptitudes à écouter et à parler. Lors de l'analyse des sous-corpus à l'aide du *Lexical Frequency Profile* (Laufer et Nation, 1995), l'augmentation prévue quant à l'emploi de mots moins fréquents n'a pas été notée. D'autres analyses effectuées à l'aide de trois méthodes accessibles sur le site Web [www.lexutor.ca](http://www.lexutor.ca) (indice des mots gréco-latins apparentés, décompte des familles de mots et rapport des types par famille) ont montré que même si les apprenants ont continué d'employer dans une grande proportion des mots fréquents, leur vocabulaire productif comprenait moins de mots apparentés à des mots français, une plus grande variété de mots fréquents et un plus grand nombre de formes développées sur le plan morphologique. L'article parle

aussi des conséquences de ces résultats sur la recherche en acquisition du vocabulaire basée sur la fréquence et sur l'enseignement du vocabulaire.

Consider the two stories below written by Marie-Eve, a young franco-phone learner of English in Quebec. The first text was written at a point when this 11-year-old beginner had experienced 100 hours of intensive ESL instruction. The second text was written at the end of the program, six months later, after approximately 400 hours of intensive instruction. Both texts were written in response to pictures; the prompt for the first text depicted the discovery of some new kittens, while the second showed a schoolyard altercation.

#### Text 1 (100 hours)

My grandmother is sick. Why? I don't no. She have begin to sick the night of November 13. We haven't assez (= *enough*) of dollars to help my grandmother. I'm very sad. It's not just for she and for me! I thing que (= *that*) for my cat: Lady, is sick. It not eat and is faible (= *weak*). Finaly, one week after, my cat, Lady, have four kittens, My grandmother feel good and she is very happy because I have help Lady to have her young cats.

#### Text 2 (400 hours)

My story is in a school yard. Yannick and his friends play soccer on the grass. The have fun. Laura is there too but she can't play soccer because the boys don't want a girl in their team. Laura is angry because she know that the boys lets play Yannick: You and your friends are just macho! Yannick say to Laura: You're a girl and the girl like you can't play soccer. Laura say to them: Why? Because you are not strong. You allow the other boys but not she. Laura say to can play with the Barbies but not soccer. She was so angry that she throw a rock on Jeffrey's head. After, Laura hit the stomach of Yannick with her foot. The boys are very suprised by the force of Laura that they lets she play with them. End

Although the later text still contains many errors, there are clear signs of language development. One dramatic difference is text length. The first story is only 81 words long, but 400 hours of instruction later, Marie-Eve is able to produce a coherent text that is almost twice as long (145 words). Another difference is the absence of French words in the second text; in the first, she resorted to three (*assez*, *que*, and *faible*) in order be able to complete her story. Her use of the colourful adjective *macho* suggests an increased ability to use more infrequent English

words – possibly in tandem with an increased awareness of the shared lexical resources of English and French (*macho* means ‘aggressively masculine’ in both languages).

The starting point for the research reported in this paper was the assumption that the writing of Marie-Eve and other learners like her would show evidence of increased lexical richness over time. This assumption is explored in a series of learner corpora produced by over 200 French-speaking learners of English at four points during their sixth year of schooling. These staged corpora offered us a unique opportunity to trace lexical development and to determine the appropriate tools for doing so.

### Background

The study took place in the French-speaking province of Quebec. The participants were Grade 6 students enrolled in French elementary schools. All learners were taking part in special intensive ESL programs in which they devoted close to half of their regular school year (approximately 400 hours) to the learning of English. Intensive ESL programs in Quebec differ from traditional immersion models in that they are not content-based: The regular Grade 6 curriculum (science, math, French language arts, etc.) is completed in French in a condensed format that frees up time for ESL (for descriptions of intensive ESL see Collins, Halter, Lightbown, & Spada, 1999; Spada & Lightbown, 1989). The writing samples that formed the corpora for this study were collected as part of a larger longitudinal study designed to assess the relative effectiveness of different distributions of intensive instruction (Collins, White, & Springer, 2005). Student progress was monitored at regular intervals of 100 hours of instruction, using a variety of comprehension and production measures.

The measure of lexical richness used in the study is the lexical frequency profile or Vocabprofile; this computerized analysis technique was developed and tested by Laufer and Nation (1995). Vocabprofile software determines the proportions of running words (tokens) in a submitted text that can be found in each of the following lists: a list of the 1,000 most frequent word families of English, and the 1,001–2,000 most frequent list (both by West, 1953), and Coxhead’s (2000) Academic Word List (a list of word families that occur frequently in academic writing). Words that do not occur on any of the previous frequency lists are categorized as ‘off-list.’ Results of the analysis of Marie-Eve’s second story using Cobb’s Vocabprofile software (an online version of Vocabprofile is available at [www.lex tutor.ca](http://www.lex tutor.ca)) are shown in Figure 1.

FIGURE 1

Vocabprofile output for Marie-Eve's 400-hour narrative (see Vocabprofile page at [www.lextutor.ca](http://www.lextutor.ca) for details of these measures)

	Families	Types	Tokens	%
First 500	...	...	113	77.93
<b>K1 words (1–1,000)</b>	<b>47</b>	<b>54</b>	<b>117</b>	<b>80.69</b>
Function	...	...	81	55.86
Content	...	...	36	24.83
Anglo-Sax (not Greco-Lat/Fr cog)	...	...	24	16.55
<b>K2 words (1,001–2,000)</b>	<b>7</b>	<b>7</b>	<b>8</b>	<b>5.52</b>
Anglo-Sax	...	...	6	4.14
<b>AWL words (academic)</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>0.69</b>
Anglo-Sax	...	...	1	0.69
<b>Off-list words</b>	<b>?</b>	<b>7</b>	<b>19</b>	<b>13.10</b>
Total	55+ ?	69	145	100.00

Words in text (tokens)	145
Different words (types)	69
Type–token ratio	0.48
Tokens per type	2.1
Lex density (content words/total)	0.44

Pertaining to on-list only	
Tokens	126
Types	62
Families	55
Tokens per family	2.29
Types per family	1.13
Anglo-Sax index (A-Sax tokens + functors / on-list tokens)	88.89%
Greco-Lat/Fr-Cognate Index: (Inverse of above)	11.11%

Figures in the bolded rows show that 81% of its 145 words are from the 1,000 most frequent families, 6% are from the 1,001–2,000 list, 1% is from the Academic Word List (AWL), and 13% are off-list. (In Marie-Eve's story, the off-list words were the items *soccer* and *macho* and names of characters such as *Yannick*.) Unlike measures based on type–token ratios that are suited to capturing the extent to which L2 learners use more varied lexis as their proficiency develops, Vocabprofile indicates the extent to which they use more 'difficult' (i.e., less frequent) vocabulary. In other words, rather than focusing on the growing numbers of different words learners use (which would normally be expected to increase over time), the measure is able to tell us something about the kinds of words they use (see Daller, Van Hout, & Treffers-Daller, 2003, for a discussion of this distinction). In addition to the proportions of words at each of four frequency levels, the output of the online version of Vocabprofile offers other details about the kinds of words that feature in a submitted text. As Figure 1 shows, information about the proportions of function, content, Greco-Latin-based and non-Greco-Latin-based words is available, along with counts of types, tokens, and families. However, our initial interest was in tracing learners' lexical development over time using Vocabprofile's frequency list feature.

This method of analysis offers several advantages in addition to providing pedagogically useful insights into the kinds of words learners use. First, since Vocabprofile draws on widely recognized frequency lists, it offers an objective way of defining what 'sophisticated' or 'unusual' lexis is in a way that allows for useful comparisons across a variety of learning and research contexts. Secondly, Vocabprofile has a track record of delineating lexical development in both cross-sectional (Laufer & Nation, 1995; Laufer & Paribakht, 2000); and longitudinal (Laufer, 1994, 1998) studies of learners of English. These studies found that a basic divide between proportions of frequent vocabulary in a learner production (words on lists of the 1,000 and 2,000 most frequent families) on one hand, and less frequent items (academic and off-list words) on the other, proved to be a reasonably reliable indicator of proficiency. The studies show that as learners advance, the proportion of lexis in the academic and off-list categories tends to increase.

Most of the studies have been conducted with adult university students at intermediate levels and beyond, but Laufer and Nation (1995) suggest that Vocabprofile may also be a useful way of assessing beginning learner productions. Vermeer's (2004) research provides support for the effectiveness of using lexical frequency profiling to assess young learner productions: Of the four measures of lexical richness he tested, a measure based on frequency lists derived from a

corpus of Dutch classroom input proved to be the most reliable in distinguishing between native and non-native oral productions. He argues that frequency-based approaches are particularly well suited to investigating child lexical development, given research that has found a relationship between the frequency of words in language input and vocabulary learning in children.

A previous study of 10 intensive ESL classrooms in Quebec used Vocabprofile to examine the characteristics of the classroom input that learners were exposed to (Meara, Lightbown, & Halter, 1997). The study determined that the overwhelming proportion of word types in transcripts of classroom input (about 96%) could be found on the list of the 2,000 most frequent word families of English (West, 1953); these results were highly consistent across the 10 classrooms investigated. Initially, the researchers interpreted the findings as evidence that the input the learners received was lexically impoverished; later they revised this view, speculating that input consisting largely of items at the 2,000 frequency level could be considered lexically rich, if, as they surmised, the young Quebec learners' vocabulary knowledge was very basic. Our investigation builds on this work by specifying more closely what such learners know, and how this lexical knowledge develops. More specifically, we focus on the extent to which learners in intensive classrooms are able to transform the large amounts of 2,000-level vocabulary they are exposed to into productive knowledge. In this, we took up Laufer and Nation's (1995) suggestion that the divide between 1,000-level vocabulary and words not in this category is likely to be a useful indicator of development in beginning learners' written productions. Thus, a text produced by a learner after 100 hours of instruction might contain a very high proportion of words from the list of the 1,000 most frequent word families and few from the 1,001–2,000 most frequent, AWL, and off-list categories, but the proportion of items in the later categories could be expected to increase over the course of 400 hours of intensive instruction. Our initial research question was as follows: Do texts show evidence of increased lexical richness (as indicated by decreasing proportions of 1,000-level words) over 400 hours of instruction?

## **Methodology**

### *Participants*

Approximately 230 French-speaking learners of English in nine classes participated in the research. All were 11- or 12-year-olds in intensive ESL

programs in French medium primary schools in Quebec. All schools from which participants were drawn were located in French-speaking communities located an hour outside of the city of Montreal where students have little exposure to English outside of class. Prior to joining the intensive ESL programs in year six of their schooling, the participants had received about an hour of ESL instruction per week during years four and five; therefore, they can be classified as beginning learners of English.<sup>1</sup> Participants completed a battery of language measures including the narrative-writing task at regular intervals – after approximately 100, 200, 300, and 400 hours of ESL instruction.<sup>2</sup> Participant numbers varied slightly from one test session to another, as the result of absenteeism.

#### *Learner texts*

At each of the four testing sessions, participants were presented with a picture to respond to in writing. A new picture prompt was used in each session in order to make the task more interesting for the participants, to avoid possible familiarity effects, and to provide the learners with contexts that would allow them to exploit their rapidly developing knowledge of English. Each of the four prompts featured age-appropriate subject matter and a clear focus for a narrative (see Collins, Halter, Lightbown, & Spada, 1999, for an example). Students were instructed to describe what they thought had happened before the event shown, what was being depicted at the moment, and what might happen next. Prompts in the initial stages required some knowledge of words for animals, family members, and parts of the body; later prompts presented possibilities for more complex scenarios. Pictures were chosen in consultation with Grade 6 ESL teachers, and with knowledge of the general vocabulary themes that the children encountered at different points in the program. This selection was particularly important in the early stages of the program, when students had very limited knowledge of English. They did not have access to dictionaries for this task. To encourage them to write as much as they could, they were told that they could insert a French word in their texts if they were not able to render a concept in English. They were given 20 minutes to complete the writing.

#### *Procedures*

A graduate student research assistant, a native speaker of English with extensive ESL teaching experience and a high level of proficiency in

French, transcribed the handwritten narratives exactly as they had been written, complete with spelling errors. An extract from a learner's text is shown below.

Original version of writer's text:

One day a boy took the pwesonal look at the girls and the girl was not very happy. After on hour she return see the boy and she kik her foot on his body. Afte a woman come hand she was not happy because the girl kik her foot on the boy. The girl explen the situation and the woman not chicane the boy ...

The texts were then altered in two ways. First, spelling errors of recognizable English words such as *pwesonal* and *kik* were corrected to *personal* and *kick* so that the lexical-frequency profiling software could process these items as English words and assess their use in the corpora. Without this adjustment such items would automatically be categorized as off-list along with infrequent words and proper names. The correction process sometimes required interpretation of the narrative context. For example, in the segment above, it seems clear that the intention of the participant who wrote 'hand she was not happy' was 'and she was not happy.' Grammar errors (agreement, case, etc.) were not corrected.

In addition, all French words were identified by adding the suffix tag *French*. (Learners had usually indicated the use of a French word by offsetting it in quotations.) This step allowed us to track the use of French items in the corpora, including items such as *vent* (sells) and *rayon* (ray), which the profiling software would otherwise treat as English words. The adjusted version of the text above appears below.

Edited text:

One day a boy took the personal look at the girls and the girl was not very happy. After one hour she return see the boy and she kick her foot on his body. After a woman come and she was not happy because the girl kick her foot on the boy. The girl explain the situation and the woman not chicaneFrench the boy ...

The transcribed narratives were then assembled in computer files, one for each of the four measurement intervals. Narratives produced by 20 participants who had been absent for one or more of the testing sessions were removed so that each of the four corpora consisted of the work of the same 210 participants. Two further steps were taken to prepare the corpora for Vocabprofile analysis: Following procedures established by Laufer and Nation (1995), all proper names (e.g., *Laura*,

*Yannick*, and *Barbie* in the example above) were manually removed from the texts; all words that had been tagged as French were also removed. The size of each of these corpora is shown in Table 1. The picture of increase over time suggests that the young learners became progressively more fluent in being able to write at length in their L2.

Each corpus was then submitted to lexical frequency analysis using the online Vocabprofile program. Because we were interested in group comparisons over time rather than the progress of individual learners, we treated each corpus as a single text instead of submitting individual narratives to the analysis.

### Analyses and results

To answer the question of whether the apparent fluency increases also entailed differences in ability to use more infrequent words over time, we turn to the results of the lexical frequency analysis.

#### *Lexical richness of the narratives*

The Vocabprofile output for the four corpora is shown in Table 2. The first row shows percentages of words from the list of the 1,000 most frequent families (K1). Here we see that in contrast to what was hypothesized, the proportions of words in this most basic frequency category did not decrease over time. In fact, there is an increase from 85% coverage at 100 hours to 91% at 200 hours; after this point, proportions of words in this category appear to stabilize at around 90%. It is rather surprising that proportions of words from the 1,001–2,000

TABLE 1  
Corpus size in numbers of words at four measurement points (210 writers)

100 hours	200 hours	300 hours	400 hours
15,159	18,069	21,246	22,984

TABLE 2  
Percentages of words in four frequency levels in learner narrative corpora produced at intervals of 100 hours of instruction

	% at 100 hours	% at 200 hours	% at 300 hours	% at 400 hours
K1	85.31	90.81	89.18	90.88
K2	11.47	6.22	7.56	5.23
AWL	0.26	0.35	0.30	0.42
Off-list	2.96	2.62	2.96	3.47

level (K2) – the zone where growth might be expected for these beginning learners – proportions decrease over time (with a minor disruption to the pattern at 300 hours). The off-list proportions (the category where infrequent items like *macho* can be found) shown in the last row remain stable at about 3% throughout. The proportion of items on the AWL is negligible in all four corpora. In sum, there is no indication that the narratives used increasingly more lexis from the lower-frequency levels over time; in fact, it appears that proportions decreased (K2) or remained unchanged (off-list).

To determine whether the differences over time in proportions of 1,000-level words were significantly different, we used the following procedure: Each corpus was divided into 42 shorter segments, each consisting of narratives produced by five randomly selected learners. Each segment was then submitted separately to Vocabprofile for analysis; the result was sets of coverage percentages for which means could be calculated and compared. The five-part segments were matched across the four corpora so that segment 1 of each corpus was made up of narratives produced by writers A, B, C, D, and E, segment 2 by writers F, G, H, I, and J, and so on. Each segment was then submitted separately to Vocabprofile for analysis. The rationale for dividing the corpora into segments produced by five writers was based on the need for longer texts. Laufer and Nation (1995) have determined that 200 words is the minimum length needed for reliable Vocabprofile measurement, but almost all of the individual narratives were well under this length, so treating the narratives individually was not an option. As Table 2 shows, the shortest corpus contained narratives produced after 100 hours of instruction. In this corpus, the five-part segments ranged in length from 243 to 543 words; mean length was 358 words ( $SD = 75$ ); thus all segments in the four corpora were well over the 200-word criterion. Lengths in the subsequent corpora were consistently larger (see Table 3). Experimentation with samples of varying lengths revealed that profiles of compiled segments were reasonably stable as long as they were 200 words or longer, so comparisons of texts of varying sizes seemed justified.<sup>3</sup>

TABLE 3  
Sizes of five-narrative segments in numbers of words

	100 hours	200 hours	300 hours	400 hours
Mean ( <i>SD</i> )	358 (75)	431 (78)	505 (109)	567 (105)
Range	243–543	262–623	306–837	398–807

The proportions of 1,000-level words in the productions across the four time intervals were then compared using a one-way repeated measures ANOVA. The results are displayed in Table 4. Means for proportions of 1,000-level words in the 42 segments (see the first column in the lower half of Table 4) closely resemble the proportions determined by analyzing the corpora in their entirety (see the first row of Table 2). There was a significant difference in the 1,000-level proportions over time, and post-hoc Tukey's HSD for related samples revealed that the difference was between the initial writing time (after 100 hours of instruction) and the subsequent three testing times.<sup>4</sup> Other differences also proved to be statistically significant (200 vs. 300 hours, and 300 vs. 400 hours), but the largest differences over time are between earlier (100-hour) and later (200-, 300-, and 400-hour) performance.

### Discussion of initial results

These findings show that young, beginning learners of English do not incorporate increasing amounts of less frequent vocabulary into their written productions over the course of 400 hours of intensive instruction.

However, we were hardly prepared to conclude that no lexical development had occurred over the 400 hours. Therefore we decided to re-examine the corpus data from a different perspective, looking more carefully at changes within use of frequent words, using tools available at the online Vocabprofile site ([www.lex tutor.ca](http://www.lex tutor.ca)). The decision to do so

TABLE 4  
Repeated measures ANOVA: K1 coverage

Source	SS	df	MS	F
Time	898.37	3	299.46	94.79*
Subjects	174.98	41		
Time X subjects	388.57	123	3.16	
Total	1,461.92	167		

\*  $p < 0.01$

Tukey's HSD test for related samples

M	SD	Time (hours)	100 hours	300 hours	200 hours	400 hours
85.18	2.25	100				
89.10	1.69	300	*			
90.82	1.69	200	*	*		
90.87	1.71	400	*	*		

\*  $p < 0.05$

was based on a number of characteristics of the writing that were observed during the transcription of the texts. First we noticed that the learners relied on L1 to fill lexical gaps in the L2, especially in the earlier corpora. The second indicator that we considered was the extent to which these young francophones relied on words that are cognate with French in their English productions. English contains thousands of words of Romance language origin, and Marie-Eve uses a number of them (e.g., *surprised* and *stomach*) appropriately in her story of a schoolyard fight. But there are also instances where French cognates are less suited to the register of simple stories told by 12-year-olds. For instance, one participant wrote, 'He is thinking about the *sensation* that the boy can have in his heart the next day.' A 12-year-old native speaker of English telling this story would likely speak of the boy's *feelings* instead. Such cases of using infrequent Romance-based words like *sensation* (an off-list word) in places where more frequent Anglo-Saxon words like *feelings* (1,000-level) seem more appropriate recurred in the data: Characters often *respond* rather than *answer*, *supervise* rather than *watch over*, and *reject* rather than *say no*.

Third, we realized the limitations of an approach that measures lexical development in terms of increasing use of words beyond the 1,000-most-frequent level. Changes in the ways these early beginner writers used very frequent English words might also indicate development. For instance, if the data showed that the learners incorporated an increasing variety of 1,000-level lexis into their narratives as the intensive course progressed, this would also be a clear evidence of lexical growth (i.e., increase in productive vocabulary size). We were also interested in the extent to which learners used morphological forms of frequent words in early and late productions. Here the focus is not on greater numbers of words but on ability to use known words in more complex ways. We had the impression that the later corpora contained more instances of inflected forms (e.g., *played* and *playing* vs. *play*) and also more derived forms (*punishment* vs. *punish*), so we set out to investigate this systematically across the staged corpora.

In sum, in our reappraisal of the data, we investigated the following characteristics of the staged corpora: reliance on French words, reliance on Romance-based lexis, variety of word families within the 1,000 level, and morphological variety of items within word families. We predicted that over the course of instruction, use of French itself as well as the use of French cognates would decrease. We expected that learners would also use a wider range of basic vocabulary and that there would be increased morphological variation.

### Re-examining the staged corpora

To evaluate the extent to which the writers used French words to fill gaps in narratives, we returned to earlier versions of the corpora from which French words and proper nouns had not yet been removed. We then relied on the list function of the Vocabprofile software. In addition to calculating the frequency percentages shown in the sample output in Figure 1, online Vocabprofile also provides lists of all tokens found in a particular frequency category. All French words that had been tagged were automatically placed in the list of off-list tokens (along with English words not found on the frequency lists). A sample segment of output from the 400-hour corpus is shown below:

```
charles cherchefrench cherchefrench chew chicanaitfrench chicanefrench
chicanefrench chicanefrench chicanefrench chili china chinese chinese
choquefrench chum cicanefrench cindy cindy cindy cindy circuitfrench
classroom classroom coeurfrench colerefrench colletfrench consolerfrench
contest coufrench coupfrench coupfrench coupfrench
```

Our research team printed out the lists of off-list tokens and counted numbers of tokens tagged as French for each of the eight corpora. These totals were then used to calculate proportions of the corpora that were accounted for by French words.

To assess levels of Romance cognate use, we used the cognate index available on the Vocabprofile site (based on wordlists by Bauer & Cobb, 2000) to evaluate the corpora from which French words and proper nouns had been removed. The cognate index measures the proportions of Romance-based families on the 1,000, 2,000, and AWL lists that a submitted text contains. In Figure 1, which shows the output for Marie-Eve's story, this information appears in the last line, where it can be seen that 11% of the lexis this learner used was of Romance-language (usually Greco-Latin) origin. To assess the numbers of different 1,000-level word families used in the four corpora (French and proper names removed), we consulted Vocabprofile output for families (see column 2 in Figure 1). In this instance, we were interested in numbers of families in the K1 (1,000-level) listing (shown in the second row). To determine the number of families, Vocabprofile matches words in a submitted text to lemmatized lists. These are lists made up of headwords (e.g., *believe*) along with their inflected and simple derived forms (e.g., *belief*, *beliefs*, *believed*, *believes*, *believing*, and *believer*).<sup>5</sup> Thus if two writers used the

form *believed* in a story and three used *believing*, this would register as a single family in the analysis of these five stories.

We were also interested in the extent to which morphological variants such as *belief*, *believed*, *believing*, etc., occurred in the corpora. Vocabprofile uses the lemmatized lists described above to calculate the mean number of types per family for all 1,000-, 2,000-, and AWL-level words in a submitted text.<sup>6</sup> In our hypothetical example of five stories that featured two types (*believed* and *believing*) from one family (*believe*), the software would assign a value of 2 for the family *believe*. Each of the other words that occurred in the five sentences would be evaluated in a similar fashion. The mean types-per-family figure for Marie-Eve's text can be seen in the fourth line from the bottom in the sample of her Vocabprofile output shown in Figure 1. The figure of 1.13 indicates that few of the 55 word families she used appeared in more than one of its possible inflected or derived forms.

## Results of the re-examination

### *Reliance on French*

As expected, the need for French to compensate for lexical gaps in English decreased considerably over time. Results in percentages of the total word count that the French tokens accounted for are shown in Table 5. After 100 hours of instruction nearly 1 word in 16 (~6% of the total corpus) was French; after 400 hours, this proportion has been reduced dramatically to approximately 1 word in 50 (~2%).

The kinds of French words that learners needed also changed over time. Analyses of the nature of the gaps that were filled by L1 words provide further insights into the learners' lexical development. As might be expected, the most commonly supplied French words were relatively low-frequency words needed to describe a particular feature of the picture prompt (e.g., the French translation equivalents of *pregnant* in a cat with kittens picture, *needle* in a hospital picture prompt, and *punch* in a schoolyard fight scene). An interesting exception to the finding that most insertions were topic-specific occurred in the 100-hour narratives: The most frequently supplied French word in the corpus was one with

TABLE 5  
Percentages of French words in learner corpora at four measurement points

	100 hours	200 hours	300 hours	400 hours
French words (%)	5.86	4.18	2.63	2.27

an important grammatical function, illustrated in the following sentences taken from three different learners' texts:

Amelie is very happy and *sa* grandmother too  
 The little girl return in *sa* house  
 She kiss *sa* mother.

Learners needed French in these cases to render the possessive determiner *her*.<sup>7</sup> Possessive determiners in English are an attested challenge for Romance speakers (White, 1998; White, Muñoz, & Collins, in press), and the reliance on L1 to render this relationship clearly signals a grammatical need. In addition, 70% (25 of 36) of the unknown content words rendered in French at 100 hours were verbs; most have English translation equivalents that are high-frequency (1,000-level) words of Germanic origin, such as *ask, bring/take, come/go back, give, find, keep, leave, look for, and say*.

#### *Reliance on cognates*

The French cognate indices based on Vocabprofile output are shown in the first row of Table 6. The first line in this row shows the results of the analyses of whole corpora while figures in the second line report means based on analyses of segments. The two sets of figures resemble each other closely. Analysis of the segment means using a one-way repeated measures ANOVA (and post-hoc Tukey's HSD for related samples comparisons) revealed that these changes were all significant (see Table 7). Generally, reliance on French cognates decreased over time (with a minor exception at 300 hours) – as predicted. The most pronounced difference is between the proportion accounted for by words of

TABLE 6  
 Vocabprofile output for three measures at four time intervals for entire corpora and segments (means)

Measure		100 hours	200 hours	300 hours	400 hours
Romance-based words (1,000 2,000, & AWL levels) (in %)	Corpus	18.38	12.75	13.37	11.83
	Segments	18.53	12.71	13.82	11.82
Numbers of families (1,000-level)	Corpus	431	462	473	498
	Segments	85.57	101.00	112.05	119.79
Types per family (1,000-, 2,000-, & AWL- levels)	Corpus	1.31	1.37	1.43	1.45
	Segments	1.14	1.15	1.20	1.21

TABLE 7  
Repeated measures ANOVA: Greco-Latin coverage

Source	SS	df	MS	F
Time	1,125.09	3	375.03	126.15*
Subjects	198.12	41		
Time X subjects	365.66	123	2.97	
Total	1,688.87	167		

\*  $p < 0.01$

Tukey's HSD test for related samples

M	SD	Time	400 hours	200 hours	300 hours	100 hours
11.82	1.56	400 hours				
12.71	1.30	200 hours	*			
13.82	1.84	300 hours	*	*		
18.53	2.50	100 hours	*	*	*	

\*  $p < 0.05$

Romance language origin at 100 hours and all of the subsequent measures.

#### *Numbers of 1,000-level families*

The number of 1,000-level families in the learner corpora appear in the middle row of Table 6. Analysis of mean family counts for segments using the repeated measures ANOVA and post-hoc Tukey's HSD procedures described above revealed that word-family use changed significantly over time for all pair-wise comparisons (see Table 8). As they progressed through the intensive ESL program, learners consistently used more word families within the 1,000 level to construct their narratives.

#### *Types per family*

The types-per-family indices assessed morphological variation (i.e., use of forms such as *believes*, *believing*, *believer*, etc.) and are displayed in the bottom row of Table 6. A repeated measures ANOVA (see Table 9) showed that the change over time was significant; the post-hoc Tukey's HSD indicated that the differences were between the first half of the program (the 100- and 200-hour testing times) and the second (the 300- and 400-hour times). In other words, learners' lexical knowledge developed in tandem with their grammatical knowledge, as they demonstrated an increased ability to render different forms of the same word to express different types of grammatical relationships.

TABLE 8  
Repeated measures ANOVA: Number of K1 families

Source	SS	df	MS	F
Time	27767.02	3	9255.67	67.32*
Subjects	25138.53	41		
Time X Subjects	16910.73	123	137.49	
Total	69816.28	167		

\*  $p < 0.01$

Tukey's HSD test for related samples

M	SD	Time	100 hours	200 hours	300 hours	400 hours
85.57	14.45	100 hours				
101.00	15.02	200 hours	*			
112.05	16.96	300 hours	*	*		
119.79	17.42	400 hours	*	*	*	

\*  $p < 0.05$

TABLE 9  
Repeated measures ANOVA: Type per family

Source	SS	df	MS	F
Time	.12	3	.040	37.45*
Subjects	.05	41		
Time X subjects	.13	123	.001	
Total	.30	167		

\*  $p < 0.01$

Tukey's HSD test for related samples

M	SD	Time	100 hours	200 hours	300 hours	400 hours
1.14	.03	100 hours				
1.15	.03	200 hours				
1.20	.03	300 hours	*	*		
1.21	.04	400 hours	*	*		

\*  $p < 0.05$

## Discussion

The exploration of the staged learner corpora using features of the online version of Vocabprofile identified patterns of lexical development in the young French-speaking learners of English. However, the main pattern was different from what was expected: Results over time showed that instead of ever-greater proportions of less frequent words, the learners used greater proportions of very frequent (1,000-level)

words. This distinction was most pronounced in the comparison between narratives produced before and after 200 hours of intensive ESL instruction. The proportion of 1,000-level words amounted to about 85% after 100 hours of instruction, rose to about 90% after 200 hours, and remained at this approximate level in the subsequent corpora. Not surprisingly, this is a higher proportion than occurs in texts produced by native speakers of English. According to figures reported in Nation (2001), the proportion of 1,000-level words in native speaker literary texts amounts just over 82%. Vocabprofile analysis of a graded reader simplified for intermediate learners of English (*Nothing but the Truth*, Kershaw, 1999) identified the same figure (82%) for proportions of 1,000-level words.

At first glance, the increased proportions of high-frequency words found in this study suggested a lack of progress. This impression is mistaken, as counts of French words and three other measures available through Vocabprofile revealed. Reliance on L1 vocabulary decreased over time, with the percentage of French tokens reduced to a little over 2% on the final measure. The cognate indices pointed to a reduction in use of Romance-based English words over time, from around 18% on the first measurement to around 13% on the subsequent measures. The pattern of early high use of cognates on the first measurement followed by lower levels from 200 hours onwards is a mirror image of the 1,000-level findings identified by the lexical frequency profiling, and may explain the increase in 1,000-level words over time. Evidence for this interpretation is illustrated by the sentences below, which were taken from the 100-hour corpus:

The father *search* in the forest but don't have the animal.

The grandmother with the boy *search* your cat.

Cats are not *complicated* in the house.

After an eternity of *reflection*, I called my grandma. She can help me.

All of the italicized cognates (*search*, *complicated*, and *reflection*) appear on the 2,000-level list; their non-cognate equivalents *look for*, *hard*, and *thinking* are all 1,000-level items. Thus use of cognates may account for the relatively high proportions of 2,000-level words on the first measure (and the corresponding low 1,000-level proportions). But as students' learning progressed, their productive lexicons appear to have featured more of the non-cognate 1,000-level words that are more appropriate to the narratives. That is, characters increasingly *feel* rather than have *sensations* and *answer* rather than *respond*.<sup>8</sup> The increased variety of 1,000-level words used in the staged corpora and the increased types-per-

family indices also indicate progress. Not only did the young learners use more word families over time, they also produced a wider range of inflected and derived forms of the words within these families.

The initial finding that lexical richness did not increase over time (as indicated by no apparent increase of 2,000-level+ words over time) raises the question of whether the use of a measure based on frequency lists derived from analyses of adult native-speaker corpora was appropriate in this study of child learners. While such corpora provide a baseline against which child language may be compared, frequency lists based upon a corpus representing child native-speaker language would allow for more fine-grained age-equivalent comparisons. It seems likely that lists based on adult language are an inaccurate reflection of frequent and infrequent lexis in the L2 input the participants were exposed to (Snellings, Van Gelderen, & De Gloppe, 2002); the usefulness of the AWL category seems particularly questionable, and it is not surprising that contributions to this zone were negligible (whether an age-appropriate AWL would yield different results merits future consideration). The finding that the participants used a greater variety of 1,000-level words over time hints at increases in lexical richness that more nuanced and appropriate frequency lists might have been able to identify. Evidence in support of this view comes from Vermeer's study (2004) in which he found evidence of increased use of lower-frequency lexis in more proficient child learners in his investigation of the lexical profiles of L1 and L2 Dutch productions. This research relied on frequency lists derived from a corpus of age-appropriate classroom input and distinguished between nine frequency levels.

The study made use of a new measure, the types-per-family index. To our knowledge, the index has not been used in previous research – at least not in the computerized format available through online Vocabprofile.<sup>9</sup> The finding that the measure detected an increase over time in the variety of morphological forms of words the learners were able to produce inspires confidence in its validity, especially since it occurred in tandem with another indicator (increase in number of 1,000-level families). Nicolae (2003) noted that instruments used in vocabulary research too often neglect the syntactic aspects of words; this measure may prove useful in addressing the imbalance. Further experiments to test its validity and reliability in other contexts are warranted.

It is clear that the other measures used in conjunction with Vocabprofile, including a measure of lexical variation, were needed to provide a more complete picture of the learners' progress. This finding is consistent with the combined approach advocated by Daller et al. in their 2003 study. They used a frequency-based measure in combination

with a measure of lexical variation to assess lexical richness. In the present study, the cognate index – numbers of 1,000-level families and types per family within this level – allowed us to interpret change over time among young low-level L2 learners of English. Overall, the study has provided evidence in support of the importance of the 1,000-level band for understanding L2 vocabulary development among lower-level learners. However, the 1,000-level band may be relevant for explaining aspects of more advanced learners' vocabulary as well. Edwards and Collins (2006) employ mathematical modelling tools (based on Zipf's Law, following Meara, 2005) to demonstrate that analyses of the proportions of words within the 1,000-level band yield productive vocabulary size estimates for L2 groups at a range of vocabulary levels.

### **Implications for learning and teaching**

Results based on a variety of measures showed that intensive instruction contributed to substantial changes in the young ESL learners' productive vocabulary knowledge. We have argued that the increased use of frequent vocabulary is a positive development, if, as we have seen, it goes hand in hand with using vocabulary that is also more register-appropriate, diverse, and morphologically complex. However, lexical development also implies control over more unusual words and raises the following question: How many more hours of instruction might be needed to increase the use of vocabulary beyond the 1,000-most-frequent level? Is there a role for more focused vocabulary instruction to facilitate the process? An investigation of lexical richness by Granger (1996) suggests that some pedagogical intervention may indeed be warranted. Her study of essays produced by advanced adult francophone learners and native speakers of English showed that in comparison to the native speakers, the ESL learners tended to overuse high-frequency English words of largely Germanic origin and to underuse lower-frequency words that are cognate with French. Our findings with beginner-level francophones suggest that there may be an initial stage of overuse (or at least inappropriate use) of these lower-frequency cognates to compensate for lexical gaps; but they are replaced by higher-frequency words (often of Germanic origin). While this change resulted in word choices that were more appropriate for the meanings that the young learners wished to express, Granger's study demonstrates that moving beyond these high-frequency words may present a persistent challenge, even when learners have an apparent cognate advantage.

Investigations of lexical richness among learners from different L1 backgrounds and at different levels of ESL proficiency would allow us

to obtain a more complete picture of the development of frequent, infrequent, and cognate vocabulary knowledge. Longitudinal studies of learners from various L1 backgrounds (Ortega & Ibarra-Shea, 2005) would also help chart the developmental path. There is also a need for a better understanding of how L2 learners perceive the semantic overlap between their first and second languages (Kellerman, 1986; Ringbom, 1986), and of the academic words that young L2 learners need to know to be able to read and produce age-appropriate texts. Cummins (1984) has pointed to the difficulties school-aged L2 speakers encounter when they lack sufficient proficiency in the target language to handle academic content. More research along the lines of Gardner's (2004) comparisons of the vocabulary content of different types of reading materials for children could inform the creation of an eventual AWL for young learners. Such research may be able to inform responsive, needs-based instructional interventions so that L2 learning can become more effective. With these goals in mind, we hope that the baselines for beginning-learner productions outlined in this study will serve as a useful basis for comparisons in future investigations and a foundation for improving instruction.

**Marlise Horst** is associate professor at Concordia University, where she teaches L2 vocabulary acquisition, TESL methodology, and the history of English. Her research interests include L2 vocabulary learning through extensive reading and computer-assisted vocabulary learning. She has taught English in Egypt, Saudi Arabia, Oman, Hong Kong, and North America, and holds a PhD in applied language studies from the University of Wales, Swansea. In addition to teaching and research, she enjoys quilting, travel, and cooking.

**Laura Collins** is associate professor at Concordia University, where she teaches courses in second language acquisition, pedagogical grammar, and cross-linguistic influence. She has published studies on the acquisition of tense and grammatical aspect, and the impact of different distributions of instructional time on L2 learning. Her current research projects focus on the interaction between lexical, morpho-syntactical, and phonological development in second language acquisition, and the application of insights from this research to L2 pedagogy.

### Acknowledgements

We would like to recognize the contributions of colleagues Tom Cobb and Joanna White and research associate, Randall Halter, whose expertise and insights informed the analyses and interpretation of the data. We received

invaluable assistance from graduate student researchers John Gilbert, Angela Kleanthous, Michèle Plomer, and Suzanne Springer, and helpful commentary from two anonymous *CMLR* reviewers. The research was supported by grants from TESOL International Research Foundation, Social Sciences and Humanities Council, Fonds québécois de la recherche sur la société et la culture, and Concordia University.

### Notes

- 1 A few learners were advanced speakers of English (as revealed by their performance on pre-test measures at the outset of the larger study and confirmed by their teachers); texts produced by these learners were excluded from the analyses.
- 2 Performance on the other measures is reported in Collins et al. (2005).
- 3 We recognize that this departs from the usual Vocabprofile practice of comparing texts of equal size (Laufer & Nation, 1995). As noted, we found that texts that were 200 words or longer tended to be stable. However, in a preliminary analysis, we divided the corpora into equal segments of 350 words each. Since the segments did not represent individual learner contributions equally, this did not allow for valid matched comparisons over time. Although the analysis was rejected for this reason, it is interesting to note that comparisons based on these equal segments produced patterns of results that were virtually identical to those reported here.
- 4 One-way repeated measures ANOVA tables were modelled after Hatch and Lazaraton (1991). For the formula used to calculate post-hoc Tukey comparisons, see Vasserstats at <http://faculty.vassar.edu/lowry/VassarStats.html>
- 5 The lists used by Vocabprofile to assign words to families are based on Bauer and Nation's (1993) definition of families at Level 3. At Level 3, a family consists of the base word, its inflections, and derived forms with the affixes *-able*, *-er*, *-ish*, *-less*, *-ly*, *-ness*, *-th*, *-y*, *non-*, and *un-*.
- 6 The software does not classify off-list words into families, since this would require master lists for hundreds of thousands of English words (i.e., all off-list words in the language). For this reason, a question mark appears in Vocabprofile output for numbers of families in the off-list category (see column 2 in Figure 1.)
- 7 There were also examples of *son* (his) and *ses* (their). However, the possessed entities in the texts were most frequently a female figure or a noun that is feminine in French, which probably explains why *sa* was the most used possessive determiner (in French, the possessive determiner agrees with the gender of the possessed entity).
- 8 Since the learners described a different picture each time they wrote, it was not possible to systematically trace changes in words used to express

particular meanings, such as changes in numbers of instances of *search* and *look for*, over the staged corpora.

- 9 The types-per-family ratio was devised expressly for this project by Tom Cobb.

## References

- Bauer, E., & Cobb, T. (2000). *Greco-Latin word families on the West's frequency lists and the AWL*. Unpublished manuscript, Université du Québec à Montréal.
- Bauer, L., & Nation, P. (1993). Word families. *International Journal of Lexicography*, 6, 253–279.
- Cobb, T. (2000). The compleat lexical tutor [Web site]. <http://www.lex tutor.ca>
- Collins, L., Halter, R.H., Lightbown, P.M., & Spada, N. (1999). Time and the distribution of time in second language instruction. *TESOL Quarterly*, 33, 655–680.
- Collins, L., White, J., & Springer, S. (2005, July). *The distribution of instructional time and second language learning outcomes*. Paper presented at the 14th World Congress of Applied Linguistics, Madison, WI.
- Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Clevedon: Multilingual Matters.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213–238.
- Daller, H., Van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics*, 24, 197–222.
- Edwards, R., & Collins, L. (2006, June). *Can lexical frequency profiles estimate vocabulary sizes? Some insights from applied mathematics*. Paper presented at the joint meeting of the American Association for Applied Linguistics/ Canadian Association of Applied Linguistics, Montreal.
- Gardner, D. (2004). Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics*, 25(1), 1–37.
- Granger, S. (1996). Romance words in English: From history to pedagogy. *KVHAA Konforenser*, 36, 105–121.
- Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Boston, MA: Heinle & Heinle.
- Kellerman, E. (1986). An eye for an eye: Crosslinguistic constraints on the development of the L2 lexicon. In E. Kellerman and M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 35–48). New York: Pergamon.
- Kershaw, G. (1999). *Nothing but the truth*. (Cambridge English Readers). Cambridge: Cambridge University Press.

- Laufer, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal*, 25(2), 21–33.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, 19, 255–271.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, 307–322.
- Laufer, B., & Paribakht, T.S. (2000). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48(3), 365–391.
- Meara, P. (2005). Lexical frequency profiles: A Monte Carlo analysis. *Applied Linguistics*, 26(1), 32–47.
- Meara, P., Lightbown, P.M., & Halter, R.H. (1997). Classrooms as lexical environments. *Language Teaching Research*, 1, 28–47.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nicolae, I. (2003). *Words have grammar: Investigating aspects of word knowledge for verbs*. Unpublished master's thesis, Concordia University, Montreal.
- Ortega, L., & Ibarra-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, 25, 26–45.
- Ringbom, H. (1986). Crosslinguistic influence and the foreign language learning process. In E. Kellerman and M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 150–172). New York: Pergamon.
- Snellings, P., Van Gelderen, A., & De Glopper, K. (2002). Lexical retrieval: An aspect of fluent second language production that can be enhanced. *Language Learning*, 52(4), 723–754.
- Spada, N., & Lightbown, P.M. (1989). Intensive ESL programs in Quebec primary schools. *TESL Canada Journal*, 7(1), 11–32.
- Vermeer, A. (2004). The relation between lexical richness and vocabulary size in Dutch L1 and L2 children. In P. Bogaards and B. Laufer (Eds.), *Vocabulary in a second language* (pp. 173–189). Amsterdam: Benjamins.
- West, M. (1953). *A general service list of English words*. London: Longman, Green.
- White, J. (1998). Getting the learners' attention: A typographical input enhancement study. In C. Doughty and J. Williams (Eds.), *Focus on form in classroom SLA* (pp. 85–113). Cambridge: Cambridge University Press.
- White, J., Muñoz, C., & Collins, L. (in press). The his/her challenge: Making progress in a 'regular' second language program. *Language Awareness*.