
Variability in ESL Outcomes: The Influence of Age on Arrival and Length of Residence on Achievement in High School

Hetty Roessingh

This article integrates findings from earlier research (Roessingh & Kover, 2003; Roessingh, Kover, & Watt, 2005) linking distinct patterns of achievement for diverse age-on-arrival (AOA) cohorts of ESL learners on the grade 12 Alberta English language arts (ELA) examination to their vocabulary and reading comprehension scores on a standardized measure over time. Recasting the data and conducting simple statistical procedures can offer further insights into the features of cognitive academic language proficiency (CALP): the relationship between vocabulary development and academic performance. I consider ESL program effects and the connection between age on arrival, vocabulary size, and achievement outcomes as reflected on the ELA examination. I compare the ESL students' scores with those of a random sample of their native-speaking (NS) academic counterparts to note patterns among the various cohorts of learners. The results suggest that measures of language proficiency (e.g., vocabulary) can be used to gain direct insights into students' academic achievement. This work has important implications for the development of theoretical growth models that would establish language-learning trajectories of good ESL progress for varied AOA and lengths of residence (LOR) fitted against a NS trajectory.

Cet article intègre les résultats de recherches antérieures (Roessingh & Kover, 2003; Roessingh, Kover, & Watt, 2005) qui établissent des liens entre, d'une part, des modèles distincts de performance à l'examen d'anglais en douzième année pour diverses cohortes d'âge à l'arrivée d'apprenants d'ALS et d'autre part, leurs scores sur des tests de vocabulaire et de compréhension à la lecture, au fil du temps et selon une mesure normalisée. La refonte de données et l'emploi de quelques procédures statistiques simples peuvent fournir un aperçu additionnel des caractéristiques de la compétence langagière cognitivo-académique, c'est-à-dire, le rapport entre le développement du vocabulaire et la performance académique. L'auteure a porté son attention sur les effets des programmes ALS et le lien entre l'âge à l'arrivée, la richesse lexicale et le rendement tel qu'indiqué par l'examen d'anglais en 12^e année. À la recherche de patrons parmi les différentes cohortes d'apprenants, elle a puisé dans les résultats à cet examen et a comparé les notes des apprenants en ALS à celles d'un échantillon aléatoire de leurs homologues académiques qui sont locuteurs natifs. Les résultats indiquent que des

mesures de compétence langagière (p. ex., le vocabulaire) peuvent fournir des indications du rendement scolaire des élèves. Cette recherche a des répercussions importantes pour le développement de modèles théoriques sur la croissance qui établiraient des trajectoires d'apprentissage langagier indiquant des progrès en ALS pour différentes cohortes (âge à l'arrivée et durée de résidence) contre une trajectoire pour les locuteurs natifs.

The limits of my language are the limits of my world. (Wittgenstein, 1889-1951)

Introduction

This study builds on a suite of previously published work (Roessingh & Kover, 2002, 2003; Roessingh, Kover, & Watt, 2005) that looked at grade 12 English language arts achievement for diverse age-on-arrival (AOA) cohorts of English as a second language (ESL) learners. The findings of these studies were highly counterintuitive. Most strikingly, we found a consistent pattern of older-arriving students (aged 15-16) outperforming the younger arrivals (aged 12-14). The youngest arrivals (aged 6-11) struggled the most in developing cognitive academic language proficiency (CALP) despite their apparent advantage in having acquired native-like proficiency with the phonological and syntactic features of English and the assumed advantage that a greater length of residence (LOR) in Canada might accrue in acculturating to community and school life. We¹ have often been challenged to explain and interpret these findings with greater specificity for both ESL and mainstream practitioners who on a daily basis must grapple with placement and program decisions for ESL learners of various AOA and LOR in Canada.

Although the scores of all age cohorts converged toward 60% on the Alberta grade 12 English language arts (ELA) diploma examination, it was in Part A, the written-response component of the examination, that distinct patterns emerged among the six subscores (see Table 2). Part B consists of multiple-choice questions and is not of direct relevance to this study. The Part A subscore patterns provided the catalyst for the study. In the current study, we examine the patterns in greater detail for insights into the role that vocabulary plays as an underlying variable in determining success on the subscores of the written-response component, as well as the overall score on the exam. The broad question that frames this study is: What is the relationship between language (measured through a standardized instrument, in this case the Gates MacGinitie Reading Tests) and educational achievement as reflected on the grade 12 (academic stream) ELA examination for varied age cohorts of ESL learners?² More specific questions include:

1. What is the relationship between vocabulary and reading comprehension scores on the Gates MacGinitie Reading Tests for ESL learners of different AOA?
2. What is the relationship of the *Choice* subscore (largely a vocabulary measure) of the ELA examination to the other subscores?
3. What is the best predictor of overall achievement on the ELA diploma examination?
4. Can we gain any insights from comparing the results of Cohorts A (no ESL support) and B (late ESL support) as to ESL program effects? (750 hours of direct contact over 3 years).
5. How do ESL students' overall scoring profiles compare with those of their native-speaking (NS) counterparts?

We begin this article with a theoretical framework for the study and a brief review of the related literature. We then explain the study design and methodology and present and discuss the findings. Finally, we suggest directions for future research. Our goal is to tease out of the complexity of second-language development the key variable(s) that can account for and identify thresholds for academic success for ESL learners.

Theoretical Framework

Our work is located at the nexus of a language-and-thought metaphor and its role in conveying meaning and having and using language. Finally, in the field of second-language acquisition, we consider the role of common underlying proficiency (CUP) and the related variables of critical mass, AOA, and LOR. The research literature in each area is briefly reviewed below.

Language and cognition. The research literature in the field of first-language acquisition and cognitive development has long recognized the connection between language and cognition (Vygotsky, 1978; Piaget, 1967; Eisner, 1997). This is also a question that philosophers ponder (Richter, 2006, re Wittgenstein; Polanyi, 1983). Although language and cognition are clearly not the same, they are closely intertwined in a complex reciprocal relationship: language is the crucial tool for shaping and enhancing mental frameworks for making sense of the world; and at the level of symbolic and abstract thought, language is critical to the process of mediating meaning. Thus language may be thought of as an instrument or tool that in its application or use, allows for communicative exchange, thought, and academic school work for children.

Recently research interest has increased in establishing the link between language—especially vocabulary measures—and cognition. Several recent studies (August, Carlo, Dressler, & Snow, 2005; Marzano, 2004; Hart & Risley, 2003) suggest that size of vocabulary is one of the best single measures of overall intelligence and the ability to think and indeed may act as a proxy for cognition, especially in early childhood (to age 9).

We have only to reflect on the words of Helen Keller (1902): they are a poignant reminder of the potential of language to transform our ability to make sense of the world and, moreover, the effect of our actions and consequent feelings for our actions on this world as these shape not only conscious awareness, but also conscience.

Suddenly I felt a misty consciousness as of something forgotten—a thrill of returning thought; and somehow the mystery of language was revealed to me. I knew then that “w-a-t-e-r” meant the wonderful cool something that was flowing over my hand. That living word awakened my soul, gave it light, hope, joy, set it free! ... Everything had a name, and each name gave birth to a new thought.... On entering the door I remembered the doll I had broken. I picked up the pieces. I tried vainly to put them together. Then my eyes filled with tears; for I realized what I had done, and for the first time I felt repentance and sorrow. (p. 34)

In the second-language acquisition literature, Cummins’ (1981) framework for the development of basic interpersonal communication skills (BICS) and cognitive academic language proficiency (CALP) is widely recognized and at a macro level provides a useful conceptual model for our work here. Most practitioners are familiar with the premise that language and context are interrelated and develop along a continuum from less to more cognitively demanding tasks, and from high levels of contextual supports to low. The fourth quadrant of Cummins’ models is characterized by abstract thought, symbolic uses of language, and especially metaphor.

The role of metaphor and culture. Lakoff and Johnson (1980) make a compelling argument for the pervasiveness of metaphor in constructing meaning, particularly for complex and abstract constructs. Thus, for example, aging, illness, and time are universal ideas that are often framed and understood by way of metaphors such as seasons (e.g., *the autumn years of life*), war (e.g., *a courageous battle with cancer*), money (e.g., *invest some time*):

Many of the concepts that are important to us are either abstract or not clearly delineated in our experience (the emotions, ideas, time, etc.) ... We need to get a grasp on them by means of other concepts that we understand in clearer terms. (p. 115)

Although the transcendent universality may be recognized across cultures, metaphor may be constrained or reflected by culture, making it difficult for second-language learners to interpret the meaning. Peterson and Coltrane (2003) contend that “students cannot truly master a language until they have also mastered the cultural contexts in which the language occurs” (p. 1).

Research evidence is emerging that the development of metaphoric competence for second-language learners is far more complex than one might think (Weitzel & Davidson-Shivers, 2004; Boers, 2000). Although the surface,

or visible, features of culture such as food, clothing, and certain conventions of polite behavior, for example, might be understood by the cultural spectator or tourist, this is only the tip of the iceberg (Hanley, 1999). The vast amount of cultural information one needs to acquire to live and be of the culture is located below the level of the iceberg, rendering it largely invisible to the casual observer. In closing a business transaction, for example, often what is not said is more important than what is said (Campbell, 2007).

The question remains: Can culture be effectively taught, or must one be of the culture in order to acquire the culture, because culture by definition is shared by in-group members and is largely the unconscious, invisible, privileged, insider information that is the everyday taken-for-granted (Delpit, 1988; Hanley, 1999)? Older arrivals, who have already developed metaphorical competence in their first language, would seem to be at an advantage: many metaphors convey universalities that can be recognized across cultural and linguistic boundaries. Younger arrivals, who have not reached the level of symbolic thought in either their first or second language and who are unable to transfer cultural information may, therefore, be especially challenged in interpreting metaphor without instructed support. For the second generation of ESL learners (those referred to in the literature as Canadian-born or Generation 1.5), the challenge of acquiring sufficient cultural grounding in their native language, which they are learning in a minority context in Canada, is surely overwhelming.

Having and using language. As noted above, having or possessing the tools of language and using them to achieve certain academic tasks are two separate concepts that are not dichotomous, but rather interdependent phenomena: competence and performance (Chomsky, 1957); *langue et parole* (de Saussure, 1959); explicit and implicit knowledge of language and the ability to mobilize linguistic knowledge in the performance of a task (pragmatics). The present study considers this interdependence of *having* and *using* by searching for the relationship between vocabulary, reading comprehension, and the completion of academic writing tasks as required on the grade 12 ELA examination. If we can establish this relationship, we can move ahead in making the link between English-language proficiency and successful educational achievement. This strategy is less fraught than the attempt to develop descriptors of language use alone as an organizational rubric for benchmarking educational progress and achievement because this is essentially a qualitative question, far more subjective and open to individual teachers' biases. This is reflected in the discrepancy between English teachers' school-based marks and those achieved on the departmental examinations, especially for ESL learners (Roessingh & Kover, 2003). Establishing thresholds for academic work by using standard measures as an approach finds support among many researchers (Saville Troike, 1991; Marzano, 2004), and it may be a welcome complement to the present-day focus in many jurisdictions on

establishing benchmarks in the K-12 realm in Canada (Manitoba Education, Citizenship and Training, 2006; Alberta Education, 2006), the United States (as a consequence of the *No Child Left Behind* legislation), and farther afield in Australia, New Zealand, and the United Kingdom.

Common underlying proficiency, L1, L2 and AOA, LOR: The critical mix. Common underlying proficiency theory (CUP) posits that underlying knowledge structures are fluid and can transfer between languages (Cummins, 1981). Students who have already developed a high level of mathematics knowledge, for example, do not have to relearn this information: they must map new language onto preexisting mental models. This is a far less onerous task than constructing new knowledge in a language that the learner does not yet have under control.

It follows, then, that older arrivals—students who have already developed linguistic and cognitive maturity in their first language—would be at an advantage in engaging with the curriculum demands of school (Twyford, 1987/1988). At a fairly low level of proficiency in English, they might access CUP and enjoy academic success. Various researchers have identified this lower threshold at approximately a grade equivalent (GE) 4 (Lee & Shallert, 1997; Collier, 1987/1988, 1992, 1995). Many older arrivals (i.e., aged 15+) have begun their studies of English in an English-as-a-foreign language (EFL) context before immigrating and arrive with this lower threshold in place. The unanswered question in the literature relates to the higher threshold required to compete academically on a par with their NS peers and how quickly diverse AOA cohorts can achieve this threshold. Can they beat the academic clock by their grade 12 year when many ESL students aspire to write the mandated examinations and continue on into tertiary educational settings?

For elementary-aged (5-11) and junior-high-aged (12-14) arrivals, we ask the same question: What linguistic threshold in English must be in place to achieve academic success at the grade 12 level? The dynamic and shifting interaction between L1 and L2, CUP and LOR will determine this outcome. It is not fixed, but rather varies for various cohort groups. Not all marks of 60 are equal.

An important distinction among the cohorts is that of additive versus subtractive bilinguals. Older arrivals, as is clear from the above discussion, are learning their second language through the lens of the first: fully developed at age-appropriate levels, with all the additional benefits of advanced curriculum knowledge in mathematics and sciences; metacognitive awareness and strategic competence, cultural capital, strong identity, and disciplined study habits. Although younger arrivals might have the edge in sounding good, they are nevertheless forever chasing a moving target as they struggle to catch up in English at the same time as they lose their first language. Most, it seems, never do close this gap. Those who arrive between

the ages of 12 and 14 (i.e., junior high school) face other learning challenges. They may be in linguistic limbo for an extended period with neither language developing sufficient strength for the cognitive demands of high school. They reflect a profile that is neither additive (older) nor subtractive bilingual (increasingly referred to in the literature as Generation 1.5). Collier's (1987/1988, 1992, 1995) studies place the 12-14-year-old arrivals at the highest risk for educational failure of all AOA groups.

Study Design

Participants

Our research takes place in an academic high school in a large urban school jurisdiction that is increasingly characterized by linguistic diversity: approximately 21% of the students across the jurisdiction are coded for ESL support. Table 1 summarizes the learner profile of five cohorts of successful high school graduates from our school. Cohorts A through D are ESL learners. Between January 1998 and June 2002, 55 ESL learners had graduated from our school. On graduation, and in compliance with local school board and university ethics board requirements, students individually provided written informed consent to use their high school transcript data, Gates MacGinitie Reading Test scores, and relevant demographic information from their cumulative records for the purposes of this study. Cohort E consists of a random sample of NS taken from the general population at the school who may be considered to be the ESL students' academic peers.

Grade 12 English language arts examination data. The grade 12 ELA examination consists of two parts: Part A, the written response, and Part B, the reading section comprising 70 multiple-choice items. Each part is worth 50% of the total exam mark, and this mark is worth 50% of the total mark for the grade 12 English language arts course (the school-based mark provided by the English teacher is worth the other 50% of the final course mark). In our study, we focus on the results of Part A because this provides deeper insights into how students can use language to convey understanding and meaning.

Part A consists of two assignments: The Reader's Response to Literature (considered the minor assignment) and the Literature Composition (considered the major assignment). The Reader's Response to Literature provides an opportunity for students as readers to respond to a selection in the context of a given topic. Students may explore ideas informally and by referring to personal experience; however, responses must be focused on the given selection and topic. In January 2000, for example, students were required to focus on Boland's (1999) poem "The Necessity for Irony." The Literature Composition assignment requires students to use literature that they have studied to discuss the significance of a person's perspective. They may expand or

Table 1
 Profile of Four Cohorts of ESL Learners (Successful Graduates) and a
 Comparison Group of Native Speakers (NS), 1997-2002

<i>Cohort</i>	<i>Immigrant Class</i>	<i>Age Arrival</i>	<i>Risk Profile</i>
A N=4	Business	Elementary 6-11	Little or no ESL support. Academic expectations but marginal performance throughout high school in English literature courses. Approximate reading equivalent 5 on entry into grade 10. Failure and dropout of program at university level.
B N=9	Business	Elementary 6-11	Little or no ESL support. Academic expectations. At risk for academic failure in high school. Approximate reading GE 5 on entry into grade 10. ESL support offered to this cohort for grades 10-12.
C N=20	Business	Junior High 12-14	Varying amounts and types of ESL support. Academic expectations. At risk for academic failure in high school. Approximate reading GE 5 on entry into grade 10. ESL support offered throughout grades 10-12.
D N=15	Business	Senior High 15-17	New arrivals to Canada. Academically competent, in a hurry to go to university. Approximate reading GE 5 on arrival. ESL program designed to support academic goals.
E N=31	N/A: Native speakers	Canadian-born	High academic achievers. Expected to attend university, fulfilling entrance requirements of a minimum overall average of 76%—generally, the lowest average admitted to the local university in those years.

augment the idea introduced in the first writing assignment (described above).

The ELA diploma exam is offered several times each year, responding to the needs of students whose programs are semestered (January and June sittings); who follow the quarter system (November and April sittings); or who attend summer school. On completion, the examinations are sealed and delivered to Edmonton, where they are marked by teams of ELA teachers trained and experienced for this undertaking. The results are scrutinized and carefully monitored by Alberta Education, ensuring consistency and reliability from sitting to sitting of the examination.

The results for Part A are most meaningful in the context of the examination blueprint and the scoring descriptors. The blueprint is included in the Appendix. The blueprint, the scoring descriptors, and a thorough analysis of the examination results are available on the Examination Manager's Report

(renamed Assessment Highlights in 2004) that is published after each sitting of the examination and can be found on the Ministry's (Alberta Education) Web site. The six scoring categories (subscores) are displayed in Table 2 for the four cohorts of ESL learners (by AOA) and a comparison cohort of native speakers' results. They include: Thought and Detail for Reader's Response (TD/RR), Writing Skills for Reader's Response (WS/RR), Thought and Detail for Literature Composition (TD/LC), Organization for Literature Composition (Org/LC), Matters of Choice—essentially, a reflection of vocabulary (Ch/LC)—and Matters of Correctness—essentially, a measure of mechanics and grammar use (Corr/LC). Each subscore is weighted, and for ease of understanding these data, we have calculated the average for each cohort and converted them to percentages. Note that it is the Thought and Detail score on the Literature Composition (TD/LC) that is the crucial score: it is weighted most heavily on the exam.

Gates MacGinitie Reading Tests—time series data. The Gates MacGinitie Reading Tests were administered at four regular intervals throughout the duration of the ESL students' time in high school. A fifth reading score was extrapolated from these results to coincide with the timing of the grade 12 ELA examination. Table 3 displays these results.

Methodology

We examined the descriptive data, looking for surface meaning and patterns; we then developed a strategy for a more detailed statistical analysis of the data that would reveal correlations among the elements in the dataset suggesting answers to the questions that framed the inquiry. Figure 1 illustrates our overall analysis strategy in working with the data.

Table 2
English Examination Results for Four Cohorts of ESL Learners
and Three Cohorts of NS, Part A.

Cohort	<i>TD</i> <i>RR/7.5</i>	<i>WS</i> <i>RR/7.5</i>	<i>TD</i> <i>LC/12.5</i>	<i>Org.</i> <i>LC/7.5</i>	<i>Ch.</i> <i>LC/7.5</i>	<i>Corr.</i> <i>LC/7.5</i>	<i>Total: 50</i> <i>Part A</i>
A: elementary no support <i>n</i> =4	50%	35%	40%	45%	35%	35%	40.0%
B: elementary with support <i>n</i> =9	67.8%	51.1%	68.89%	67.78%	58.89%	55.56%	61.78%
C: junior high with support <i>n</i> =20	62%	49.46%	57.6%	60.5%	52.5%	48%	55.28%
D: senior high with support <i>n</i> =15	60.61%	51.33%	66%	64%	56.67%	49.33%	58.8%
E 1: NS, average over 80%, <i>n</i> =10	89%	92%	87%	91.91%	93%	93%	91.0%
2: NS, average 70-79%, <i>n</i> =11	79%	80%	69%	70.9%	76.36%	80.9%	75.3%
3: NS, average 60-69%, <i>n</i> =10	72%	67%	66%	64%	68%	71%	67.8%

Note. Raw scores have been weighted and converted to percentages.

Table 3
Cohort Averages for Gates MacGinitie Reading Scores

	<i>Cohort B</i>	<i>Cohort C</i>	<i>Cohort D</i>
Gates 1: Vocab	5.6	4.9	4.0
Rd'g. Comp.	5.9	6.19	4.88
Total	5.6	5.37	4.28
Gates 2: Vocab	5.7	5.81	4.67
Rd'g. Comp.	8.87	7.9	6.21
Total	6.97	6.58	5.24
Gates 3: Vocab	6.77	6.26	5.3
Rd'g Comp.	11.67	8.43	6.98
Total	8.3	6.93	5.93
Gates 4: Vocab	8.8	7.66	6.1
Rd'g. comp.	10.1	10.4	8.85
Total	9.3	8.66	7.08
Gates 5: Vocab	8.75	8.4	7.5
Rd.'g. Comp.	10.7	10.8	9.9
Total	9.75	9.6	8.7

Pearson's r^3 was used as the measure of correlation. The relatively small sample size made it impossible in a number of instances to examine each cohort individually. In calculating correlation coefficients, it proved necessary to combine cohorts A and B (elementary with and without support) in order to get a large enough sample size.

Findings

We begin by considering the descriptive data presented in Tables 2 and 3. Table 2 displays the outcomes (by subscores) of the ELA examination. Several findings are worth highlighting.

We note that Cohorts B, C, and D (those who received instructed support) achieved the target of 55-60% on the examination. On the surface at least, it appears that they have all been successful and received comparable results. The real story requires deeper examination and interpretation to understand that some remain at educational risk whereas others are still on a fairly dynamic path leading into their academic future at university. Not all 60s are the same as it turns out. We take up this point in the discussion section.

We observe an increase on the overall score from 40% to 61.78% between Cohorts A (elementary arrivals who had no ESL support) and B (elementary arrivals who received late support throughout high school). Further, there is an increase in the *Choice* subscore from 35% to nearly 59%, with a concomitant increase in the *Thought and Detail* subscore from 40% to 68.89%.

We note the same striking pattern of lower *Choice* scores compared with *Thought and Detail* scores for all ESL cohorts. This is not the case for NS where

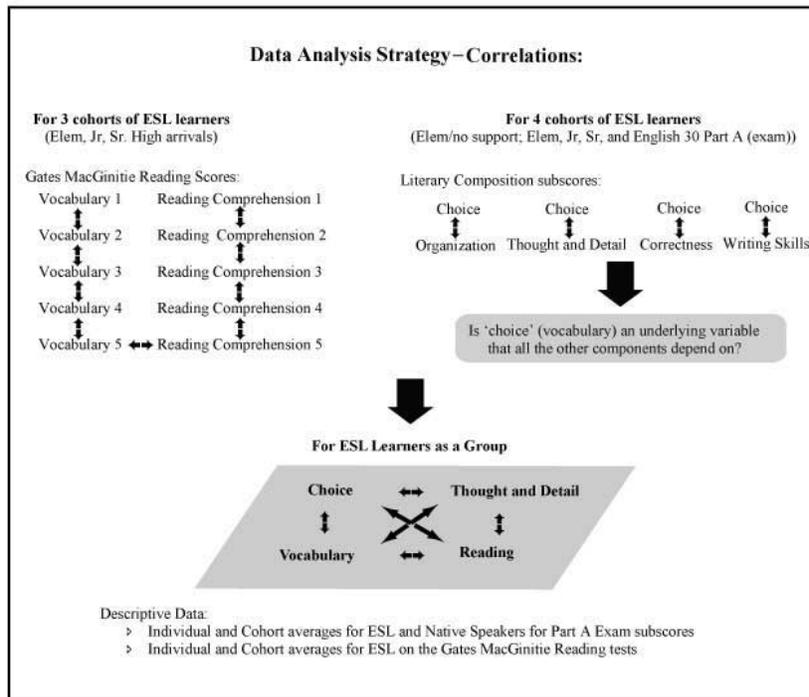


Figure 1. Data analysis strategy.

the opposite pattern holds: we see that the *Choice* score for these students is consistently higher than the *Thought and Detail* score; moreover, the gap between these two scores is small.

Turning to the Gates MacGinitie Reading scores, we see that on entry into the ESL program (and placement at the upper intermediate level), the total score for all cohorts of students is GE 4-5. For Cohorts B and D, the vocabulary and reading comprehension scores are tightly aligned, whereas Cohort C shows a discrepancy between these two scores: the reading comprehension measure (6.19) is well in advance of the vocabulary measure (4.9). Cohorts B and D begin to accelerate in reading comprehension once they have reached a vocabulary threshold: the senior high arrivals (similar to the junior high arrivals) reaching this point earlier (GE 4.7) than the elementary-aged arrivals (GE 5.7). Not only does it appear that the vocabulary threshold required to trigger reading comprehension is higher for the elementary arrivals, but the data also indicate that these students find it difficult to sustain the pace of vocabulary development that older arrivals (junior and senior high) achieve, although in the long run the youngest arrivals (Cohort B) have reached a higher vocabulary threshold. In the discussion section, we posit that the level of first-language development may help to explain these find-

ings: older arrivals have recourse to a reservoir of common underlying proficiency that younger arrivals do not. Thus although the elementary arrivals have a higher vocabulary and overall total Gates score by time 5, they may remain at educational risk at university.

For all ESL cohorts, reading comprehension scores outstrip vocabulary scores, ultimately reflecting a gap of GE 2-2.5 by the time they reach grade 12. All cohorts' reading comprehension, measuring around GE 10-11, still lags behind that of their academic NS-speaking counterparts who may be assumed to be reading well above grade at this point. Recall that Cohort D (senior high arrivals) may still be trading on their strength in their L1 and thus may expect future academic success even given this lag in their reading scores as long as they are prepared to continue to develop their L1 through independent reading.

We turn now to the results of the Pearson correlation measures among the variables shown in Figure 1. Both *Choice* and *Thought and Detail* were substantially correlated with success on the written section of the exam. Figure 2 shows the relationship between *Choice* and the score on the rest of the written section of the exam for the entire group ($n=48$). Pearson's r was .593, significant at the 0.01 level.

Figure 2 shows a number of outliers, which we decided to leave in. Examination results can vary for any number of reasons beyond the control of educators such as test anxiety, lack of sleep, or illness. To exclude the outliers would be to exaggerate the precision of the analysis.

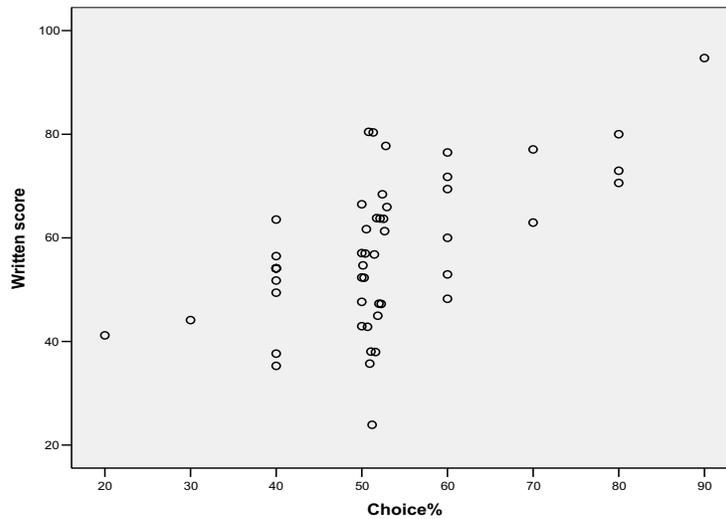


Figure 2. Choice by written score (excluding Choice score).

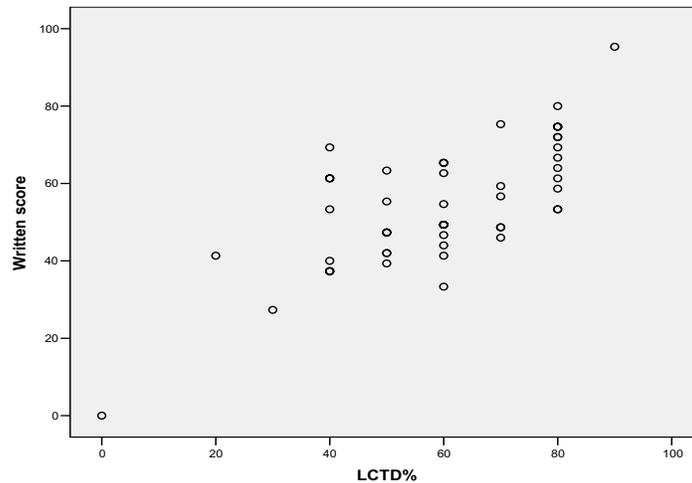


Figure 3. *Thought and detail by written score (excluding thought and detail score).*

Analysis of this relationship for each cohort showed some variability. Pearson's r indices for elementary arrivals and for senior high arrivals were respectively .734 and .666, both significant at the 0.01 level. For junior high arrivals, there was no statistically significant relationship between the two variables. We take up this point in the discussion section.

Correlation between *Thought and Detail* and the rest of Part A of the exam was substantial, with a Pearson's r of .725, significant at the 0.01 level. Figure 3 illustrates the relationship.

This correlation was similar across all three cohorts, with a Pearson's r of .711 for the elementary arrivals, .691 for the junior high arrivals, and .795 for the senior high arrivals (all significant at the 0.01 level).

Correlation between *Choice* and *Thought and Detail* for the entire group was substantial, with a Pearson's r of .779 (significant at the 0.01 level). This varied across cohorts, with no statistically significant relationship for the elementary arrivals, a Pearson's r of .793 for the junior high arrivals, and a Pearson's r of .861 for the senior high arrivals (significant at the 0.01 level). This poses intriguing riddles. For the youngest (elementary) arrivals: how are they making meaning? And for the junior high arrivals: how much longer can they ride out their CUP?

Discussion

Returning to the questions that guided this study, we discuss our findings. Our overarching question relates to language and educational achievement for various cohort groups. It would seem that vocabulary size plays a crucial role for all cohorts in the development of reading comprehension, and this in turn bears directly on the achievement outcomes. This factor is emphasized by other researchers (August et al., 2005; Moudraia, 2001; Biemiller & Slonim, 2001). However, the vocabulary thresholds required for academic success as reflected on the ELA examination varied by cohort group (i.e., age on arrival): older arrivals (Cohort D) were able to do well even with a relatively small English vocabulary (estimated at approximately 25,000 words compared with their NS counterparts who might be expected to have a vocabulary of 80,000+ words). It would seem that older arrivals are strategic in transferring meaning from their L1 with less English vocabulary than one might think. This threshold is also identified in the literature by others (Hazenberg & Hulstijn, 1996).

What is the relationship between vocabulary and reading comprehension scores on the Gates MacGinitie Reading Tests for learners of different AOA?

Table 3 displays the vocabulary and reading comprehension scores for Cohorts B, C, and D. Two findings are noteworthy in this discussion. First, students do not accelerate their reading comprehension score until after a threshold or critical mass in vocabulary has been accrued. This differs somewhat from cohort to cohort group, but in a nutshell, vocabulary GE 5-6 is required to trigger reading comprehension. This pattern is similar to that found among NS. Second, at GE 5-6, vocabulary levels begin to lag behind reading comprehension scores. The resultant gap may be as much as 2-2.5 GE. This lag is visible for all cohorts of ESL students. It may be assumed, at least for junior high and senior high arrivals, that this gap can be mediated via the underlying proficiency. The intriguing riddle is how elementary arrivals are able to construct meaning with such a sizeable deficit of vocabulary. No ready explanation is available; however, anecdotal evidence suggests that bilingual learners are able at times to make meaning directly from experience, bypassing the need to encode in language the sense they make of these experiences. This might explain the rapid accelerated growth reflected in Cohort B once they were offered ESL support in high school, often for the first time. The youngest of these students arrived in his grade 1 year. For NS, the connection between vocabulary and reading comprehension is tight: comprehension is closely aligned to the words these learners have at their disposal (MacGinitie & MacGinitie, 1992).

What is the relationship of the Choice subscore (largely a vocabulary measure) of the ELA examination to the other subscores?

It would seem that *Choice* is an underlying factor for the other subscores. Again we return to the centrality of vocabulary in virtually all areas of academic endeavor: thinking, organizing, and writing.

What is the best predictor of overall achievement on the ELA examination?

Our research evidence points to AOA and the concomitant level of L1 proficiency that we assume from the children of business- and professional-class immigrants (i.e., high socioeconomic status) as the determining variables. These students have the level of CUP required to accelerate their L2 proficiency as needed for successful achievement on the diploma examination.

Can we gain any insights from comparing the results of Cohorts A (no ESL support) and B (late ESL support) as to ESL program effects? (750 hours of direct contact over 3 years).

Instructed support has a tangible effect on all cohorts, but this effect is most visible in the achievement outcomes of the youngest (elementary-aged arrivals) reflected in Cohort B. Although it is heartening to note this accelerated trajectory, intervention should have been offered at an earlier stage in their schooling experiences in Canada. Had ESL programming been available, we believe these learners would have beaten the academic clock, crossing the vocabulary threshold needed to compete with their NS peers by high school graduation time. We were pleasantly surprised with the idea of *never too late*: these students' cognitive engines apparently had been on idle, simply waiting for the input that would trigger accelerated cognitive growth. We discussed these findings in an earlier publication (Roessingh & Kover, 2002).

How do ESL students' overall scoring profiles compare to those of their NS counterparts?

ESL learners are competing with NS students whose overall academic achievement is over 80%. The minimum requirement at many universities across Canada is a high school average of well over 80%. NS students' scoring profile is fairly even for the ELA examination. Again, the level of vocabulary required to score the *Choice* points is high, and this correlates with the ability to achieve on the *Thought and Detail* measure for these students. While ESL students' score profiles are more uneven or jagged, they can compete only if they can avail themselves of a high threshold of vocabulary that must be in place in either L1 or L2. The threshold suggested is 50,000 words or approximately 12,000 word families (Hazenberg & Hulstijn, 1996). Clearly again, the older arrivals have the advantage.

Perhaps the single most important finding in our work relates to the dynamism between the ESL learners' two underlying language systems that accounts for their academic achievement on the English diploma examination. For older arrivals (Cohort D), a mark of 59% can be achieved with the

lowest measure of English vocabulary of all cohort groups. This suggests that they have access to underlying proficiency that has already been constructed in their first language, and they are able to achieve adequate sense-making in terms of the academic demands on them in English with less vocabulary than younger arrivals. In other words, *less is more* for the older arrivals.

This is not the case for the youngest (elementary-aged) arrivals. A mark of 62% fully presses their linguistic resources into the service of making meaning. *What you see is what you get*, and for youngest arrivals who must compete with NS for a seat at university and for successful retention in their programs of study, 62% will be inadequate: NS speakers typically enter their first year with an average of 84% or better. Although the youngest arrivals showed a strong acceleration in their language development and scored the highest of all cohorts on the vocabulary measure, this would seem to have been a case of too little too late. Their academic engines had been idling for too long. We hypothesize that if we had met them and offered direct ESL support earlier, they would have been able to beat the academic clock with enough momentum to carry them forward into their tertiary studies.

Most at risk may be those who arrive between the ages of 12 and 14. They appear to be stuck between languages, with neither L1 nor L2 providing enough of what it is necessary for them to achieve academically before time runs out. In a case of *pay now or pay later*, a booster year around the age of 14 (after grade 9) might be a wise program consideration. We recognize that this may not be a palatable option for the students or their parents, especially given their relatively good (ESL) academic achievement—though marginal performance in comparison with their NS peers—on entry into high school. It may be that the lack of correlation noted for this cohort between the *Choice* and the *Thought and Detail* measures reflects this competition between two languages, with neither winning the academic edge.

Conclusion

Instructed ESL support makes a difference. A clear focus on vocabulary development produces a tangible effect on the learning outcomes of all cohorts of ESL students. A distinct dynamism underlies ESL learners' academic achievement that can be accounted for by the AOA and the assumed level of proficiency learners have in their first language as this interacts with their development of the second. *Older is better, and less is more*. This finding is well documented and echoed in the second-language acquisition literature (Collier, 1987/1988; Cummins, 1981; Hakuta, Butler, & Witt, 2000; Roberts, 1994; Lee & Shallert, 1997; Jiang & Kuehn, 2001). These findings need to be taken into account in the development of a conceptual framework for setting language benchmarks for academic success.

The findings suggest that having and using language are correlated. Vocabulary measures in this case have been a strong predictor of actually

deploying these resources in efforts to demonstrate the ability to make and convey meaning on the ELA examination. The link between vocabulary and academic achievement is also supported in the SLA literature (Saville Troike, 1991; Hart & Risley, 2003). Once again, a strong vocabulary focus in teaching produces distinct benefits for all AOA.

The next steps in our research endeavor take us in the direction of developing models for academic achievement for varied AOA and LOR that make the link between language proficiency as reflected in vocabulary measures and academic success. Genesee, Lindholm-Leary, Saunders, and Christian (2005) note that the longitudinal data required for such an undertaking are lacking. Indeed a project of this scope would require systematic data collection over at least 13 years, following children through their entire K-12 educational experiences. Further, we would need to account for varied learner profiles: What about children who arrive as refugees and who are not literate in their first language due to a lack of opportunities in their homeland? What about children who acquire their first language as language-minority speakers in Canada compared with their counterparts who immigrated? Even children as young as 5 who come from these two language-learning contexts may have different learning possibilities. Canadian-born students are increasingly reflected in the numbers of children coded as ESL or English language learners (ELL) in major school jurisdictions across Canada, and preliminary studies indicate that these children do not achieve academically as strongly as their foreign-born counterparts (Alberta Education, 2006). Our demographic landscape is shifting to include Generation 1.5 children in ways not heretofore predicted. Our research agenda continues to expand and unfold in front of us.

Postscript

Since submitting this manuscript for review, I have been able to track the participants of this study and approximately 100 more ESL graduates from local high schools through their university experiences, where new understandings now emerge. This pilot study suggests that the unfolding demands of the university curriculum far outstrip the ability of any AOA cohort group to acquire English-language proficiency rapidly enough to keep up with their NS peers. The elementary-aged arrivals never did catch up, the senior high arrivals have used up their CUP advantage and are challenged by their second year, and the demands of tertiary-level study have also taken their toll on the junior high arrivals. All AOA cohorts have struggled, as reflected in low GPAs and disproportionate course failures and repeats. These data have relevance for policy, programming, pedagogy, and service provision from K-12, as well as at the university level. But this is the topic of a future separate report.

Notes

¹I use *we* throughout this report because the analysis explores aspects of research carried out by a team including P. Kover, D. Watt, and myself.

²The Gates MacGinitie Reading Tests were chosen because of this instrument's widespread acceptance in the field as a research tool, its ease of administration and interpretation, and the multiple formats available for accumulating time series data for all age cohorts.

³Pearson's *r* is a measure of the correlation between variables. A Pearson's *r* of 1.00 would indicate a one-to-one positive correlation. A Pearson's *r* over .6 is considered to indicate a substantial relationship.

Acknowledgment

The author gratefully acknowledges partial support for this study by SSHRC grant no. 410-2006-2530.

The Author

Hetty Roessingh is a long-time ESL practitioner in the K-12 school system. In 2000 she joined the Faculty of Education, University of Calgary, where she teaches courses in methods, materials and curriculum design in the MEd TESL program. Her research interests include longitudinal studies for rates of acquisition, academic thresholds, and academic success in high school and tertiary settings. Her recent interests include determining the baseline at the onset of the educational trajectory of 5-year-old children.

References

- Alberta Education. (2006). *A review of 1-12 ESL education in Alberta*. Retrieved January 25, 2007, from: http://www.education.gov.ab.ca/ipr/eslreview/Review_of_ESL_report.pdf
- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research and Practice*, 20(1), 50-57.
- Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*, 93, 498-520.
- Boers, F. (2000). Metaphor awareness and vocabulary retention. *Applied Linguistics*, 21, 553-571.
- Boland, E. (1999). *The necessity for irony*. Retrieved June 22, 2006, from <http://www.runningland.com/2005/06/>
- Campbell, A.M. (2007). How to lose your shirt in China: The economy beckons, but its business climate makes China a killing field for Canadian firms. *Macleans*, January 29, pp. 26-29.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Collier, V. (1987/1988). The effect of age on acquisition of a second language for school. *National Clearinghouse for Bilingual Education, Focus Number 2, Winter 1987/1988*. Retrieved July 24, 2008, from: <http://www.ncela.gwu.edu/pubs/classics/focus/02aage.htm>
- Collier, V. (1992). A synthesis of studies examining long-term language minority student data on academic achievement. *Bilingual Research Journal*, 16, 187-212.
- Collier, V. (1995). Acquiring a second language for school. *National Clearinghouse for Bilingual Education, Directions in Language and Education*, 1(4), 3-12.
- Cummins, J. (1981). *Bilingualism and minority language children*. Toronto, ON: OISE Press.
- Delpit, L. (1988). The silence dialogue: Power and pedagogy in educating other people's children. *Harvard Educational Review*, 58, 280-298.
- de Saussure, F. (1959). *Course in general linguistics*. (C. Bally & A. Sechejaye, Eds., W. Baskin, Trans.). New York: McGraw-Hill (Original work published 1916).

- Eisner, E.W. (1997). Cognition and representation: A way to pursue the American dream? *Phi Delta Kappan*, 78, 348-353.
- Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2005). English language learners in U.S. schools: An overview of research findings. *Journal of Education for Students Placed at Risk*, 10, 363-385.
- Hakuta, K., Butler, Y. & Witt, D. (2000). *How long does it take English learners to attain proficiency?* (ERIC Document Retrieval Service No. ED 443275)
- Hanley, J. (1999). Beyond the tip of the iceberg: Five stages toward cultural competence. *Reaching Today's Youth*, 3(2), 9-12. Retrieved 25 July, 2008, from: <http://www.cyc-net.org/reference/refs-culturalcompetence.html>
- Hart, B., & Risley, T. (2003). The early catastrophe: The 10 million word gap. *American Educator*, 27(1), 4-9.
- Hazenberg, S., & Hulstijn, J.H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, 17, 145-163.
- Jiang, B., & Kuehn, P. (2001). Transfer in the academic language development of post-secondary ESL students. *Bilingual Research Journal*, 25, 417-436.
- Keller, H. (1902). *The story of my life*. New York: Dell.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.
- Lee, J.W., & Schallert, D.L. (1997). The relative contribution of L2 language proficiency and L1 reading ability to L2 reading performance: A test of the threshold hypothesis in an EFL context. *TESOL Quarterly*, 31, 713-739.
- MacGinitie, W., & MacGinitie, R. (1992). *Gates MacGinitie Reading Tests* (2nd ed.) *Teacher's manual*. Scarborough, ON: Nelson Canada.
- Manitoba Education, Citizenship and Training. (2006). English as an additional language (EAL) and literacy, academics and language (LAL). Kindergarten to Grade 12. *Manitoba curriculum framework of outcomes*, draft document. Retrieved July 25, 2008, from: <http://www.edu.gov.mb.ca/k12/cur/diversity/eal/framework/index.html>
- Marzano, R. (2004). *Building background knowledge for academic achievement: Research on what works in schools*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Moudraia, O. (2001). Lexical approach to second language teaching. *ERIC Digest*, June 2001. Retrieved July 2, 2003, from: <http://www.cal.org/resources/digest/0102lexical.html>
- Peterson, B., & Coltrane, B. (2003). Culture in second language teaching. *ERIC Digest*, Winter, 2003. Retrieved July 25, 2008, from: http://www.cal.org/resources/digest/digest_pdfs/0309peterson.pdf
- Piaget, J. (1967). *Six psychological studies*. (D. Elkind, Ed., A. Tenzer, Trans.). New York: Random House.
- Polanyi, M. (1983). *The tacit dimension*. Gloucester, MA: Peter Smith.
- Richter, D.J. (2006). Ludwig Wittgenstein (1889-1951). *Internet Encyclopedia of Philosophy*. Retrieved July 25, 2008, from: <http://www.iep.utm.edu/w/wittgens.htm#H1>
- Roberts, C. (1994). Transferring literacy skills from L1 to L2: From theory to practice. *Journal of Education Issues of Language Minority Students*, 13, 209-221.
- Roessingh, H., & Kover, P. (2002). Working with younger arriving ESL learners in high school English: Never too late to reclaim potential. *TESL Canada Journal*, 19(2), 1-19.
- Roessingh, H., & Kover, P. (2003). Variability of ESL learners' acquisition of cognitive academic language proficiency: what can we learn from achievement measures? *TESL Canada Journal*, 21(1), 1-21.
- Roessingh, H., Kover, P., & Watt, D. (2005). Developing cognitive academic language proficiency: The journey. *TESL Canada Journal*, 23(1), 1-27.

- Saville-Troike, M. (1991). Teaching and testing for academic achievement: The role of language development. *NCBE Focus: Occasional Papers in Bilingual Education*, 4(Spring). Retrieved August 1, 2002, from: <http://www.ncela.gwu.edu/pubs/focus/focus4.htm>
- Twyford, C. (1987/1988). Age-related factors in second language acquisition. *National Clearinghouse for Bilingual Education, New Focus*, 2 (Winter). Retrieved December 16, 2006, from:<http://www.ncela.gwu.edu/pubs/classics/focus/02bage.htm>
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. (V. John-Steiner, S. Scribner, & E. Souberman, Eds., M. Cole, Trans.). Cambridge, MA: Harvard University Press.
- Weitzel, M., & Davidson-Shivers, G. (2004). Effects of metaphors for Asian and majority-culture students. *Home Health Care Management and Practice*, 17(1), 14-21.

Appendix

Blueprint for Part A, from Examination Manager's Report, English, 30 January 2000 <http://education.alberta.ca/admin/testing.aspx>

(Note. in 2004-2005 this report was renamed *Assessment Highlights*).

Part A: Written Response

The blueprint below outlines the requirements for each assignment, the categories for scoring each assignment, the amount each category contributes to the total mark (Parts A and B combined), and the percentage of students achieving at the various levels. In January 2000, the average raw score for Part A was 32.6 out of 50.

The Part A results are most meaningful in the context of the assignments and the scoring descriptors. The most useful starting place for reviewing these results is at the **3 Satisfactory** level. Such work exceeds the pass mark of 505. The scoring guides that describe scoring criteria are in the 1999-2000 English 30 Information Bulletin located at <http://ednet.edc.gov.ab.ca/studenteval/>

Part A: Examination Blueprint and Percentage Distribution of Score

Description of the writing assignment	Scoring category	Proportion of total mark (%)	Excellent		Proficient		Satisfactory		Limited		Poor	
			5	4.5	4	3.5	3	2.5	2	1.5	1	Ins*
Reader's Response to Literature:												
The student is required to respond personally and thoughtfully to literature, and to communicate effectively in writing.	Thought & Detail	7.5	4.2	5.1	20.2	15.5	32.2	10.8	9.5	1.3	.9	.2
	Writing Skills	7.5	4.2	4.6	22.1	16.3	36.6	8.4	6.3	.8	.6	.2
Literature composition:												
The major assignment maintains a thematic connection to the minor assignment. The student is required to demonstrate an appreciation of literary works studied in class by discussing theme and the literary techniques that the author uses to support that theme. The students is also required to synthesize thoughts clearly and correctly in writing	Thought and Detail	12.5	5.7	3.9	18.6	12.7	27.3	13.5	14.6	2.1	1.3	.2
	Organization	7.5	5.1	4.2	18.1	14.8	33.6	12.9	9.4	1.0	.6	.2
	Matters of Choice	7.5	6.1	4.1	19.8	13.7	38.7	9.8	6.6	.6	.5	.2
	Matters of Correctness	7.5	6.4	4.8	21.0	14.0	33.4	10.3	8.2	1.1	.7	.2

Note. The shaded portion represents the percentage of students who achieved or exceeded the acceptable standard. Students with scores of 4, 4.5, and/or 5 have achieved the *standard of excellence*.

*Ins (Insufficient) is a special category that includes students who did not attempt the assignment, wrote too little to evaluate, or wrote answers that did not address the topic.