Applied Psycholinguistics **29** (2008), 41–65
Printed in the United States of America
DOI: 10.1017/S014271640808003X

Lexical acquisition over time in minority first language children learning English as a second language

HEATHER GOLBERG and JOHANNE PARADIS
University of Alberta

MARTHA CRAGO
Université de Montréal

Received: January 6, 2006 Accepted for publication: December 15, 2006

ADDRESS FOR CORRESPONDENCE
Johanne Paradis, Department of Linguistics, University of Alberta, Edmonton, AB T6G 2E7, Canada. E-mail: johanne.paradis@ualberta.ca

ABSTRACT
The English second language development of 19 children (mean age at outset = 5 years, 4 months) from various first language backgrounds was examined every 6 months for 2 years, using spontaneous language sampling, parental questionnaires, and a standardized receptive vocabulary test. Results showed that the children's mean mental age equivalency and standard scores on the Peabody Picture Vocabulary Test—Third Edition nearly met native-speaker expectations after an average of 34 months of exposure to English, a faster rate of development than has been reported in some other research. Children displayed the phenomenon of general all-purpose verbs through overextension of the semantically flexible verb *do*, an indicator of having to stretch their lexical resources for the communicative context. Regarding sources of individual differences, older age of second language onset and higher levels of mother's education were associated with faster growth in children's English lexical development, and nonverbal intelligence showed some limited influence on vocabulary outcomes; however, English use in the home had no consistent effects on vocabulary development.

How minority first language (L1) children acquire verbal academic skills in their second language (L2) has been the subject of much research, but comparatively less research has been focused on L2 oral language proficiency (Cummins, 2000; Saunders & O'Brien, in press). Vocabulary accumulation is a cornerstone of acquiring the majority L2 for minority L1 children. Having a vocabulary of a sufficient size is crucial for developing age-appropriate oral proficiency in the L2, as well as for literacy skills, both of which are important for educational success overall (August, Carlo, Dressler, & Snow, 2005; Verhallen & Schoonen, 1993; Vermeer, 2001). With respect to catching up to same-aged native speakers in lexical

© 2008 Cambridge University Press 0142-7164/08 $15.00

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

acquisition, one challenge L2 children face is that they are trying to hit a moving target (Cummins, 2000). What this means is that as L2 children are trying to catch up in vocabulary size, native speakers are increasing their vocabulary as well. This contrasts with the acquisition of inflectional morphology. For example, English native-speaker 6-year-olds are at ceiling in their performance with the past tense [-ed] (Rice & Wexler, 2001), and thus, in terms of catching up to native speakers, 6-year-old English L2 learners are trying to hit a stable target. The purpose of this study was to investigate how long it takes for children to hit the moving target of native-speaker vocabulary size in their L2, what developmental processes they display in their lexical acquisition, and what factors predict individual differences in their rate of acquisition.

As part of a large-scale study of the Spanish and English development of bilingual school-aged children in Miami, Cobo-Lewis, Pearson, Eilers, and Umbel (2002) examined receptive and productive vocabulary size in both bilinguals and monolinguals at various grade levels, which we refer to as the Miami Study. Bilingual children included those who spoke English and Spanish at home and those who spoke Spanish only at home (effectively English second language [ESL] children), divided into groups attending bilingual versus English-only programs. Children were also divided into low and high socioeconomic status (SES) groups. Cobo-Lewis et al. (2002) found that vocabulary development in the ESL children was slow in terms of reaching monolingual norms, which would be expected given the "moving target" nature of vocabulary accumulation; for example, even among the fifth grade groups there were differences between English monolinguals and the ESL children (see also August et al., 2005; Umbel, Pearson, Fernández, & Oller, 1992; Vermeer, 2001). Languages spoken in the home, instructional program, and SES played a role in determining outcomes; high SES children who spoke both Spanish and English at home attending English-only programs performed closest to monolinguals. The scale of the Miami Study permitted the researchers to make some important generalizations about English L2 achievements, but the design was cross-sectional, and a longitudinal examination of ESL children's vocabulary development would provide additional insights into this process, with respect to sources of individual variation in particular. Furthermore, it would be important to compare findings from Spanish-speaking ESL children in Miami to ESL children whose L1s were more minority in status, meaning not spoken extensively outside children's homes. Accordingly, the present study examined the lexical development of a group of ESL children over time. These children had a diverse set of L1 backgrounds, and resided in a community where none of their L1s was widely used.

Measuring L2 lexical acquisition need not be limited to overall vocabulary size, but could also include measuring children's growing understanding of word semantics. For example, L2 children have been shown to lag behind native-speaker age peers in their abilities to produce antonyms, synonyms, generate analogies and definitions, and understand polysemy and homonymy (August et al., 2005; Cobo-Lewis et al., 2002; Verhallen & Schoonen, 1993; Vermeer, 2001; but see Peña, Bedore, & Rapazzo, 2003). The ability to understand semantic relationships between words depends in part on children's knowledge of individual word meanings in the L2, although the latter has received comparatively less attention.

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

Among word classes, research on L1 acquisition has indicated that meanings of verbs may be more complex to grasp than meanings of nouns early on because nouns more often than verbs have stable and concrete referents, which makes it easier to map phonological form to meaning (Gleitman, Cassidy, Napa, Papafragou, & Tueswell, 2005; Tomasello, 2003). With respect to L2 acquisition of verb semantics, Harley and King (1989) and Harley (1992) found that English L1–French L2 learners in immersion programs showed a reliance on general all-purpose (GAP) verbs in their oral expression, presumably to stretch their meager verb lexical resources to meet task demands. In other words, the L2 children used high frequency verbs with nonspecific semantics to describe events that native speakers used less frequent and more precise verbs to describe. For example, to describe an alarm clock that pinched the nose to wake up a sleeper, L2 learners tended to use the more frequent and less semantically specific verb *prendre* (take/grab), whereas native speakers used *pincer* (pinch; Harley 1992). Because Harley and King's research was with immersion students who typically did not use French with their peers or outside the classroom, it is possible that this phenomenon of GAP verb reliance would be less apparent in minority L1 children, who have much more exposure to the L2 with peers at school and in the community. Standard scores from vocabulary size tests cannot tell us about verb lexical diversity in particular, and certainly cannot yield information about GAP verb usage. Therefore, alternative methods such as speech sampling would be necessary to address the question of whether the GAP verb phenomenon is characteristic of minority L1 children's L2 lexical acquisition.

It is important to point out that in prior research on GAP verb use in L1 acquisition, researchers have used both the terms GAP verb and high frequency verb to refer to this phenomenon, sometimes interchangeably (Conti-Ramsden & Jones, 1997; Rice & Bode, 1993; Thordardottir & Ellis Weismer, 2001; Watkins, Kelly, Harbers, & Hollis, 1995). However, there is a clear distinction between what constitutes a GAP verb versus a high-frequency verb. GAP verbs are verbs that are not semantically specific, and this quality permits their use in a variety of contexts, as well as their inappropriate overextension by language learners. High-frequency verbs, in contrast, are simply the verbs that are most commonly used in a language by all its speakers. When English GAP verbs are defined as verbs that do not have a semantically specific meaning, the set of GAP verbs typically consists of high-frequency verbs like *do* and *come*. The reverse is not the case, however, because not all high-frequency verbs used in English have this nonspecific semantic quality. For example, in a stretch of discourse where a speaker is retelling a past conversation, he/she might use the verb *said* very frequently, but not in such a way as to stretch a nonspecific meaning. It is reasonable to assume that L2 children would have a greater proportion of their productive vocabularies comprising high-frequency verbs than same-aged native speakers because their more frequent appearance in the input would facilitate their acquisition. However, to understand whether these high-frequency verbs are actually being used as GAPs, that is, with nonspecific semantics, individual verbs in context must be examined (cf. Harley, 1992). Therefore, this study investigated the GAP phenomenon by examining L2 children's verb use according to the semantics of the intended event being described.

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

As mentioned above, Cobo-Lewis et al. (2002) looked at two key sources of individual variation affecting L2 lexical acquisition: amount of exposure to English in the home and SES. They examined exposure to English in the home as a binary or categorical variable, that is, the families used either Spanish and English at home, or only Spanish at home. However, it is possible that this variable exerts gradient effects on L2 lexical acquisition. For example, in some minority L1 children's homes, English might be used 20% of the time and the native language 80%, and for others it might be the opposite. Previous research has suggested that gradient variations in English use in the home can influence rate of L2 development in ESL children (Jia, 2003; Jia & Aaronson, 2003). In the present study, families were categorized according to whether they were high or low English users so that we could ascertain whether this difference had an impact on children's English development.

Concerning SES, researchers have found this variable to be an influential source of variation in L1 lexical acquisition (Arriaga, Fenson, Cronan, & Pethick, 1998; Fenson et al., 1994; Hart & Risley, 1992; Hoff, 2003a; Hoff-Ginsberg, 1998). The sources of this influence were argued to be the differences in parenting style and quality and quantity of speech to children associated with differences in SES, in both monolingual American families and cross-culturally (Hart & Risley, 1992; Hoff, 2003b; Hoff, Laursen, & Tardiff, 2002). One intriguing finding from the Miami study was that SES was a significant source of individual variation in vocabulary learning in English even for children whose parents spoke only Spanish at home. Cobo-Lewis et al. (2002) commented that high SES families who reported using only Spanish at home also often reported being fluent in English, and thus, they spoke Spanish out of choice not necessity. Perhaps some English was in the home environment of these families, for example, when doing homework, even though parents reported only using Spanish with their children. In contrast, the majority of parents of the children in this study had limited English proficiency at the outset, and were new arrivals from a variety of linguistic and cultural backgrounds, that is, not all Spanish-speaking Latin American. Thus, in our study we sought to determine whether SES differences would influence L2 lexical acquisition in children from more diverse minority L1 backgrounds.

There are other potential sources of individual variation in L2 children's lexical acquisition aside from language spoken in the home and SES. For example, age of onset of L2 acquisition may be a factor. It is possible that if children begin acquiring English after kindergarten age, they might accumulate vocabulary faster because they have more cognitive and literacy-based resources for language learning that can enhance their L2 development (cf. Cummins, 1991, 2000), and because older children, whether bilingual or monolingual, display stronger lexical processing skills than younger children (Windsor & Kohnert, 2004). A middle-childhood onset has been shown to be advantageous with respect to academic verbal skills in minority L1 children (Collier, 1987, 1989), although middle-childhood onset has been found to diminish ultimate attainment in the L2 for pronunciation and morphosyntactic knowledge (Bialystok & Miller, 1999; Flege, 1999; Weber-Fox & Neville, 2001). The impact of variation in age of onset within early childhood on L2 oral vocabulary development is largely unknown. Accordingly, we compared

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

the lexical development of the younger children who began learning English before kindergarten with the older children who began learning English in or after kindergarten, keeping the amount of exposure to English constant. In so doing, we wanted to understand the role of age differences in the children's L2 lexical acquisition.

Another potential factor underlying individual differences in L2 acquisition is language aptitude, an inherent ability consisting of several analytic and working memory skills important for acquiring language structures and words. Research among adolescent and adult L2 populations has shown language aptitude to be one of the most reliable factors explaining individual differences in L2 success (Dörnyei & Skehan, 2003; Sawyer & Ranta, 2001). Although there are standardized tests, such as the Modern Languages Aptitude Test (Carroll & Sapon, 1959), available to measure this ability in adolescents and adult L2 learners, there are no standardized tests available for testing young children's language aptitude. Therefore in our study, we used nonverbal intelligence as an approximate measure of children's language aptitude, as it is known to be a related skill (Dörnyei & Skehan, 2003; Sawyer & Ranta, 2001), and Genesee and Hamayan (1980) found nonverbal reasoning skills to be predictive of individual differences in kindergarten children's L2 acquisition of French.

RESEARCH QUESTIONS AND PREDICTIONS

Concerning receptive vocabulary development, we asked how long does it take for ESL children to display similar levels of vocabulary size as same-aged native speakers? We predicted that these ESL children, as a group, would approach native-speaker norms for receptive vocabulary size more rapidly than the children in the Miami study because they would have had more limited opportunities to use their L1 outside the home; hence, more exposure and practice with English. Cobo-Lewis et al. (2002) found substantial differences between ESL and monolingual children in the second grade group, where the ESL children had had 3 years' exposure to English. We expected that by the end of our study, after nearly 3 years' exposure, these ESL children would come closer to the native-speaker norms for vocabulary size than the ESL children in the Miami study.

Regarding expressive vocabulary size, this was measured by lexical diversity in spontaneous speech rather than by a standardized test in our study. Nevertheless, we used published findings from a study of L1 acquisition as the basis of comparison (Watkins et al., 1995), and asked whether ESL children would meet native-speaker expectations for lexical diversity in production based on months of exposure to English. Given that the ESL children were older than the L1 learners, and thus more cognitively and linguistically mature, we predicted that they would meet or surpass native-speaker expectations when amount of exposure formed the basis of comparison.

Turning to the GAP verb phenomenon, we asked whether minority L1–majority L2 children would show this developmental process like majority L1–minority L2 children in French immersion. It was predicted that these ESL children would initially display overextended uses of the high-frequency, semantically flexible verb *do*, but that these GAP uses of *do* would diminish toward the end of the

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

study, in contrast with immersion students. Harley (1992) found that the GAP phenomenon persisted for as long as 5 years of schooling for immersion students; however, the L1 minority children in this study were receiving more input in the target language because they had native-speaker peers in the classroom, and the community outside the classroom was majority English speaking.

The last research question concerned the sources of individual variation for lexical acquisition in ESL children. We asked if mother's level of education, use of English in the home, age of onset of English learning, and nonverbal IQ would influence the children's receptive and expressive vocabulary growth. Based on prior research findings, we predicted that each of these variables would impact on the children's lexical development in their L2 such that higher levels of mother's education, higher use of English in the home, older age of L2 onset, and higher nonverbal IQ would be associated with faster growth in lexical development.

METHOD

Participants

The participants in this study consisted of 19 minority L1 children residing in Edmonton, an English majority language city in western Canada with a population of about 1 million. At the first interval of data collection, participants had a mean age of 5 years, 4 months (5;4, range = 4;2–6;9), and had received an average of 9 months' exposure to English in the context of a school or preschool.[1] Fourteen of the children were recent immigrants to Canada and 5 were born in Canada. Despite country of origin, parental report indicated that all the children were exposed nearly exclusively to the minority language of the home and had not received a significant amount of exposure to English until their entry into a preschool or school program in Edmonton. It is important to make clear that the five participants born in Canada had the same language history as those children who were recent immigrants in that the age of onset of English began with starting school. The participants for the study were recruited from various agencies providing assistance to recently arrived immigrant or refugee families, and through government-sponsored English language training classes for adult newcomers. Parental report also indicated that the children displayed typical language development in their L1. As determined by the Columbia Mental Maturity Scales (Burgemeister, Hollander Blum, & Lorge, 1972), all of the children had a nonverbal IQ above 85. Table 1 provides the means and standard deviations of the participants' ages and their months of exposure to English (MOE) at each round of the study. The various L1 backgrounds are reported in the footnote below Table 1.

Procedures

There were five rounds of data collection in this study, one approximately every 6 months. Children were visited in their homes two or three times during each round of testing, and they participated in a battery of language tasks. Data from two tasks, the Peabody Picture Vocabulary Test—Third Edition (PPVT-III; Dunn & Dunn, 1997) and the spontaneous speech sample, will be reported in this study

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

Table 1. *Participants' ages, months of exposure to English, receptive and productive vocabulary measures for all five rounds, and first languages*

	Round 1	Round 2	Round 3	Round 4	Round 5
Age (months)	64 (9.8)	70 (9.9)	76 (10.2)	82 (10.2)	88 (9.9)
MOE	9 (4.5)	15 (4.4)	21 (4.4)	28 (7.1)	34 (4.6)
Receptive Vocabulary (Mean Scores)					
PPVT-III					
Raw	45 (19.2)	62 (21.8)	70 (20.7)	79 (25.5)	92 (25.5)
Standard	78 (14.9)	86 (14.7)	88 (14.1)	89 (16.6)	97 (19.5)
Age equivalent	43 (15.7)	56 (18.5)	65 (18.2)	71 (21.4)	85 (25.7)
Productive Vocabulary (Mean Frequencies)					
NDW types	109 (32.4)	130 (30.0)	149 (31.5)	155 (28.1)	147 (34.0)
GAP do uses	1.6 (2.8)	1.5 (2.3)	0.7 (0.94)	0.2 (0.54)	0.4 (0.68)

Note: First language backgrounds included Korean, Mandarin, Cantonese, Spanish, Romanian, Arabic, Japanese, and Farsi. Values are means (standard deviations). MOE, months of exposure to English; NDW types, number of different word types/100 utterances; GAP do, number of uses of lexical *do* with general or overextended semantics.

(see Paradis, 2005, for information on other tasks). The Columbia Mental Maturity Scales were administered on the first visit only. Research assistants who collected the data were native speakers of English, and were either honors undergraduate or master's level students in the Department of Linguistics at the University of Alberta.

A questionnaire was administered to the parents at each round of data collection where information about parental employment and education, as well as language use in the home was gathered. An interpreter was often present to assist in administering the questionnaire to prevent any miscommunication between the parents and the research assistant. For this study, we used maternal level of education in years to determine a family's SES, as did Cobo-Lewis et al. (2002). Because these were new arrival families, the parents' relative income and occupation status in their home country was often quite different from what it was in Canada at the time of testing, and thus, level of education was deemed to be the better measure of SES in terms of determining what effect this might have on parenting style and speech to children (see also Jia & Aaronson, 2003). Language use in the home was documented using 5-point scales (1 = *native language only*, 5 = *English only*) The father's, mother's, and sibling's language use with the child, and the child's language use with each of these family members, were recorded on separate 5-point scales. The values from each scale were added together and divided by the number of scales to calculate the value for English use in the home.[2]

The data collected from the questionnaire were used to derive the values for three of the variables predicted to affect L2 acquisition over time: age of onset of exposure to English (AOE), mother's level of education (MLOE), and English

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

Table 2. *Cutoff points, means, standard deviations, and sample size for each group within the predictor variables*

Variable	Cutoff Points	Mean	*N*
AOE	≥60 months	66 (4.9)	7
	<60 months	48 (6.3)	12
MLOE	>12 years	16 (1.4)	8
	≤12 years	11 (2.3)	11
ENGHOME	≥0.5	0.58 (0.089)	11
	<0.5	0.38 (0.076)	8
NVIQ	>115	124 (8.1)	6
	85–115	104 (7.0)	13

Note: AOE, age of exposure to English; MLOE, Mother's level of education; ENGHOME, English use in the home; NVIQ, nonverbal IQ. Values are means (standard deviations).

use in the home (ENGHOME). The standard scores from the Columbia Mental Maturity Scales were used for the last predictor variable, nonverbal IQ (NVIQ). Each participant was assigned to one of two groups for each predictor variable, as determined by a specific cutoff point among the values. Table 2 shows the means and number of participants in each group within the predictor variables. The first variable, AOE, was divided based on 60 months, or 5;0 years, as the cutoff point. Therefore, those children who began to learn English at the age of 60 months or greater made up one group, and those who began to learn English at an age younger than 60 months made up the other. Sixty months was used as the cutoff point because that is the age when children begin kindergarten in Edmonton. Note that the younger L2 onset group had a mean chronological age of 58 months ($SD = 6.3$), whereas the older L2 onset group had a mean chronological age of 73 months ($SD = 6.8$), at Round 1. This means that younger/older age of L2 onset and younger/older chronological age were combined in these data, and we consider this fact in our discussion of the effect of age on the children's vocabulary acquisition. The second predictor variable, MLOE, was divided based on 12 years of education, which is the distinction between secondary versus postsecondary education. Therefore, one group contained children of mothers who had greater than 12 years of education, and the other contained children of mothers with 12 years of education or less. The third variable, ENGHOME, was divided based on a value of .5, where those families who had a value of .5 or higher made up one group, and those who had a value less than .5 made up the other. This value was a natural cutoff point as it divided the participants into to nearly equal groups numerically. Finally, NVIQ was divided according to average range (85–115) and above average standard scores on the Columbia Mental Maturity Scales. Those with scores less than 115 made up the average group, and those with scores above 115 made up the above-average group.

Spontaneous speech data were collected using a semistructured interview followed by free play in a 45-min videotaped session between the child and the research assistant. The interview consisted of questions about the child's family,

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

friends, school, and recent past and future events in the child's life. The videotapes were later transcribed and analyzed using the conventions of the CHAT and CLAN system (MacWhinney, 2000; http//www.childes.psy.cmu.edu). Ten percent of the spontaneous speech corpus at each round was independently transcribed by a different research assistant. Interrater agreement rates comparing the second assistant's versions with the originals were 91–98% for words in the transcription. The expressive vocabulary measure for this study was lexical diversity in spontaneous speech. More specifically, we measured lexical diversity as the number of different word (NDW) types in the last 100 utterances of the children's transcript. Previous studies examining lexical diversity in children with specific language impairment (SLI) have also used unique word types in a set stretch of discourse (Rice & Bode, 1993; Thordardottir & Ellis Weismer, 2001; Watkins et al., 1995). Nouns with plural or possessive morphemes, and verbs with regular past tense or progressive morphemes were counted as one word type. Contractions such as *they're* were broken down into their separate words respectively and counted as two different word types.

For the analysis of the GAP verb phenomenon in the spontaneous speech transcripts, it was decided that contextualized use of the lexical verb *do* would be examined over time. This verb was chosen for the following reasons. First, it has been cited as a GAP verb in previous studies on either English or French (*faire* [do]) learners, and thus seemed a likely candidate for semantic overextension (Harley 1992; Rice & Bode, 1993; Thordardottir & Weismer, 200l; Watkins et al., 1993). Second, the lexical verb *do* was one of the five most frequently used verbs across all rounds in this study. We examined all contexts where lexical *do* was used in the children's transcripts, coded the uses as either "contextually appropriate" or "GAP," and calculated the frequency of each type of *do* usage for each child. An appropriate usage of *do* was one where there was no obvious overextension, and no other more specific verb in English sounded more natural for the context. GAP *do* usages were those showing a nonnativelike overextension, or where a more specific verb was available and/or sounded more natural in the context. Do-support used in negation and question formation and emphatic-do were excluded from this analysis. Ten percent of the participants' use of *do* as a main verb was independently coded by a different research assistant for each round. Interrater agreement comparing the second assistant's versions with the originals was 90%.

RESULTS

The following lexical measures from the ESL children are presented in Table 1: mean PPVT raw, standard and mental age equivalency scores, mean number of different word types in 100 consecutive utterances (NDW types), and mean frequency of GAP *do* uses. To assess the trends in the children's receptive and expressive vocabulary development, one-way repeated-measures analysis of variances (ANOVAs) on the PPVT raw scores and the NDW types were conducted. PPVT raw scores were chosen over standard scores because we wanted to assess trends in vocabulary growth according to amount of exposure to English, without monolingual norm referencing by chronological age. The one-way ANOVA for

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

PPVT raw scores with Round (five levels) as the within-subjects factor showed a significant linear trend, F_{linear} (1, 18) = 86.11, $p = .000$, and no other significant trends. Paired *t* test comparisons between the means at consecutive rounds showed that mean PPVT raw scores at each round were higher than those at the previous round: 45 versus 62, t (18) = −4.92, $p = .000$; 62 versus 70, t (18) = −2.73, $p = .014$; 70 versus 79, t (18) = −3.02, $p = .007$; 79 versus 92, t (18) = −3.95, $p = .001$. The one-way ANOVA for NDW types with Round (five levels) as the within-subjects factor yielded significant linear, F_{linear} (1, 18) = 30.8, $p = .000$, and quadratic trends, F_{quad} (1, 18) = 12.10, $p = .003$. Paired *t* test comparisons between the means at consecutive rounds showed that mean NDW types at each round were higher than those at the previous round from Rounds 1 to 3: 109 versus 130, t (18) = −2.74, $p = .014$; 130 versus 149, t (18) = −3.26, $p = .004$, but no differences were found between comparisons from Rounds 3 to 5: 149 versus 155, t (18) = −0.88, $p = .392$; 155 versus 147, t (18) = 1.40, $p = .178$.

Receptive and expressive vocabulary development compared with monolinguals

Analyses using two forms of monolingual norm referencing from the PPVT were conducted to determine how the ESL children's receptive vocabulary development compared to that of native speakers over time. The PPVT standard scores in Table 1 show that the ESL children's mean score at Round 2, 86, was within the range of typically developing monolingual age mates (85–115). By Round 5, the children's mean standard score was 97, close to the expected monolingual mean of 100. However, one standard deviation below the mean for the ESL children was 77, not 85. Regarding individual differences, at Round 5, 26% (5/19) of the children had scores >100, indicating that some L2 learners can surpass the native speaker mean in just under 3 years of exposure to the target language. Conversely, 16% (3/19) of the children had scores in the range for atypical monolingual learners (i.e., < 85), even after this much exposure to the target language.

To further compare the ESL children with their monolingual peers, we examined their PPVT age-equivalency scores and their chronological ages over time. Our logic for this comparison was as follows: because ESL children catch up to native speakers eventually, they must gain vocabulary faster because they have had fewer years of exposure to the language than same-aged native speakers. Given the children's standard scores, we expected this "catching up" to be observable through plotting their actual ages with their PPVT age equivalencies in months (see Figure 1). A two-way repeated-measures ANOVA was conducted with Round (5) and Age (2: chronological age and PPVT age) as within-subject factors. There were significant main effects for Round, F (4, 18) = 76.01, $p = .000$, and Age, F (1, 18) = 15.77, $p = .001$, as well as a significant Round × Age interaction, F (4, 72) = 4.52, $p = .003$. Paired *t* test comparisons between children's mean age and age-equivalency scores at each round revealed significant differences from Rounds 1 to 4, where children's chronological age was always higher than their age equivalencies on the PPVT: 64 versus 43, t (18) = 6.44, $p = .000$; 70 versus 56, t (18) = 3.90, $p = .001$; 76 versus 65, t (18) = 3.14, $p = .006$; 82 versus 71, t (18) = 2.40, $p = .027$. In contrast, no significant difference was found at Round 5: 88 versus

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

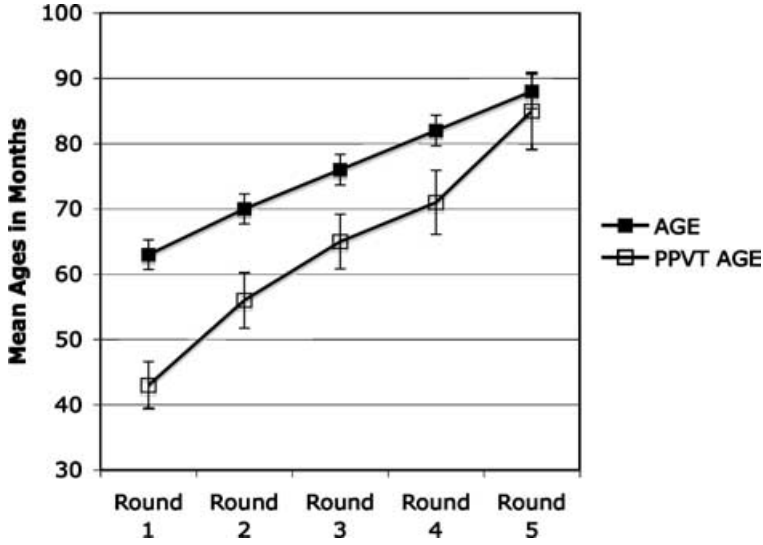


Figure 1. The mean chronological ages and Peabody Picture Vocabulary Test age equivalency scores for five rounds.

85, t (18) = 0.75, p = .462. Therefore, these data show that the children's mean chronological age was higher than their PPVT age at first, but both ages increased over time, and the trajectories intersected at Round 5. Similar to the standard scores, individual differences are noteworthy in the age-equivalency scores. At Round 5, 21% (4/19) of the children had age-equivalency scores 12 months or greater than their chronological age, and 26% (5/19) had age-equivalency scores 12 months or lower than their chronological age.

To undertake norm referencing for the measure of expressive vocabulary, we compared the ESL children's NDW types out of 100 utterances to the results published in Watkins et al. (1995) for monolingual children, because they used virtually identical methods. Watkins et al. (1995) provide NDW types for 100 utterance samples from children 39 and 59 months old. Note that for monolinguals, chronological age and months of exposure to English are identical, but not for ESL children. In Figure 2, the ESL children's NDW types are plotted at each round with chronological age and months of exposure to English marked on the x axis. The results from the monolingual children in Watkins et al.'s study are included on the basis of approximate chronological age matching (Round 1) and months of English exposure matching (Round 5). One-sample t tests revealed that when compared on chronological age, the ESL children had fewer unique word types in their samples than the monolinguals: 109 versus 160, t (18) = −6.871, p = .000, but when compared on months of exposure to English, the ESL children had more unique word types in their samples than their monolingual counterparts: 147 versus 122, t (18) = 3.215, p = .005. Unlike with the PPVT data, we do not have

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

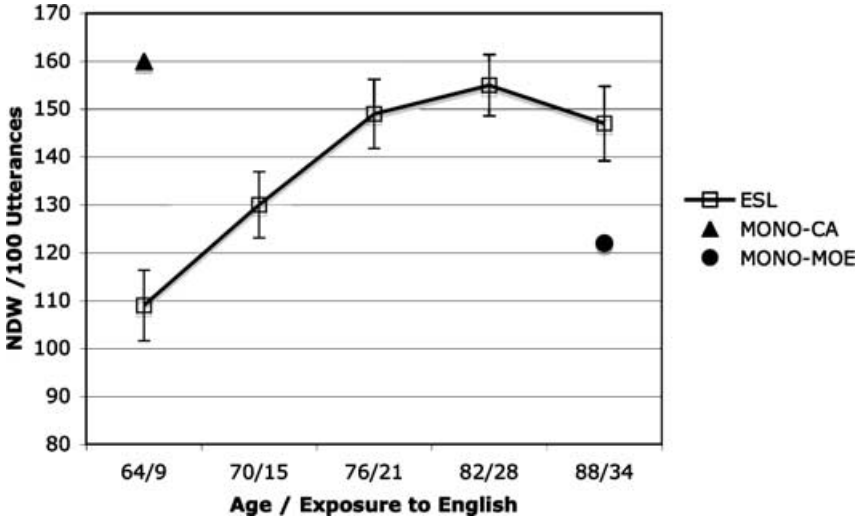


Figure 2. The mean number of different word types for five rounds for English as a second language children and monolingual age and English exposure peers from Watkins et al. (1995).

chronological age-matched monolingual data to compare at Round 5 to ascertain whether the ESL children had "caught up" to native speaker age peers.

The verb do as a GAP verb

Some examples of GAP uses of *do* are listed in Table 3 with the participant's code, their L1, and a more specific English verb deemed appropriate for the context. These examples illustrate how the occurrence of GAP *do* was spread across individuals, L1 backgrounds, and occurred even at Rounds 4 and 5. Figure 3 displays the mean frequencies of contextually appropriate and GAP uses of the verb *do* for five rounds. A two-way repeated-measures ANOVA was conducted on the *do* frequencies with Round (5) and *Do*-Type (2: context appropriate and GAP) as within-subjects factors to determine whether GAP uses of *do* diminished over time, compared with contextually appropriate uses of *do*. The analysis yielded significant main effects for *Do*-Type, $F\ (1, 18) = 6.48$, $p = .02$, G-G adjustment, and a significant Round × *Do*-Type interaction, $F\ (2.56, 46.14) = 4.65$, $p = .009$, G-G adjustment, but no significant main effect for Round, $F\ (2.91, 52.33) = 0.34$, $p = .788$, G-G adjustment. Paired t test comparisons at each round between the mean frequencies of GAP and contextually appropriate *do* yielded the following results: no significant differences at Round 1 or Round 2: 1.6 versus 1.3, $t\ (18) = -0.43$, $p = .672$; 1.5 versus 0.8, $t\ (18) = -1.13$, $p = .275$, and lower mean frequencies for GAP uses at Rounds 3 to 5: 0.7 versus 1.7, $t\ (18) = 2.21$, $p = .04$; 0.2 versus 2.3, $t\ (18) = 5.883$, $p = .000$; 0.4 versus 1.7, $t\ (18) = 4.167$, $p = .001$. These analyses reveal that GAP *do* uses were roughly equal to contextually appropriate *do* uses at Rounds 1 and 2, but that contextually appropriate *do* uses predominated from Rounds 3 to 5.

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

Table 3. *Examples of participants' utterances containing do as a GAP verb in each round*

GAP Do Utterance	Participant	L1	Appropriate Verb
	Round 1		
He *do* ribbit, ribbit	TRRK	Arabic	Say
I *do* the bigger one	SHHN	Farsi	Want
He *do* a baseball	RNL	Cantonese	Throw
How to *do* ring-around-the-rosie	RMLM	Japanese	Play, sing
One likes to *do* another castle	CHRS	Romanian	Make
	Round 2		
I can *do* any night	CNDX	Mandarin	Play
Spiderman *do* like this	THRJ	Farsi	Sprays (his web)
My dad can *do* a person	SMNS	Spanish	Draw
We *do* our name	LLKC	Arabic	Write
I *do* some grass over here too	JNNH	Mandarin	Cut
	Round 3		
Then we *did* recess	SHHN	Farsi	Have
I *did* some loud	RMLM	Japanese	Blow
[A nurse] *do* needles	LLKC	Arabic	Give
I can't *do* like craft and you know	SBST	Spanish	Make
	Rounds 4 and 5		
He *do* it like just ch . . .	THRJ	Farsi	Choke
I *do* like this (protects face with hands)	SBST	Spanish	Put my hands in front of my face
I could *do* my kites	RMLM	Japanese	Build, make

Note: L1, first language.

Sources of individual variation

Four potential sources of individual variation in vocabulary development were identified and described in the Method section: AOE, MLOE, ENGHOME, NVIQ. To determine whether there are relationships between these variables, Pearson correlational analyses were performed, and the coefficients are presented in Table 4. There was a significant negative correlation between mother's level of education and English use in the home ($r = -.531$, $p = .019$), but no other significant relationships were found.

To determine the influence of the predictor variables on children's vocabulary development, two mixed-design ANOVAs were conducted: one with PPVT raw scores and one with NDW types as the dependent variables. PPVT raw scores were chosen over standard scores for the following reason: standard scores correct for different ages in monolinguals, where higher raw scores are expected for older

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

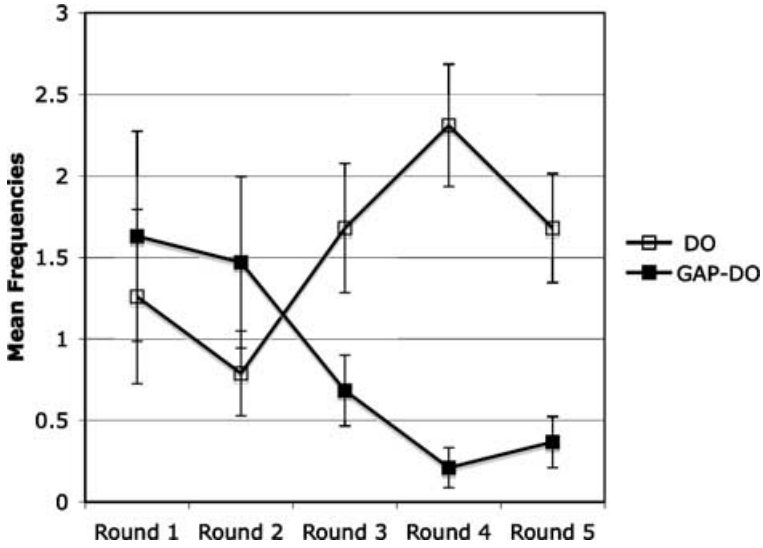


Figure 3. The mean frequencies of contextually appropriate and general all-purpose uses of *do* for five rounds.

Table 4. *Correlation coefficients between age of onset of English, mother's level of education, English use in the home, and nonverbal IQ*

	AOE	MLOE	ENGHOME	NVIQ
AOE	—	.285	−.226	.328
MLOE		—	−.531*	.343
ENGHOME			—	−.158

Note: AOE, age of exposure to English; MLOE, Mother's level of education; ENGHOME, English use in the home; NVIQ, nonverbal IQ.
*$p < .05$.

children primarily because they have had more experience learning the language. However, for ESL children, a child at 4;6 and a child at 5;6 could have the same amount of exposure to English even though their chronological ages are different. Therefore, the adjustment provided by standard scores might place false expectations on the child who is 5;6 (cf. analysis above where ESL children's PPVT scores only reached age-appropriate monolingual norms at Round 5). Hence, we reasoned that the expectations for ESL children's vocabulary size for this analysis of individual differences ought to be set on the basis of amount of exposure to English rather than chronological age. Each ANOVA had vocabulary scores as the within-subjects factor (five levels) and four between-subjects factors with two levels each: AOE (older and younger), MLOE (postsecondary and primary/secondary), ENGHOME (high and low), and NVIQ (average and above average). For PPVT

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

raw scores, the main effects of the repeated measure vocabulary scores, $F(4, 28) = 38.42$, $p = .000$, and the between-subjects factors of AOE, $F(1, 7) = 8.79$, $p = .021$, and MLOE, $F(1, 7) = 17.72$, $p = .004$, were significant, but ENGHOME, $F(1, 7) = .79$, $p = .404$, and NVIQ, $F(1, 7) = .296$, $p = .603$, were not. No significant interactions were found. Children's PPVT raw scores over time are plotted according to their between-subjects groups for AOE and MLOE in Figure 4a and 4b. The figure reveals that the means for high MLOE and older AOE were consistently higher than those for low MLOE and younger AOE. To estimate the effect size of the significant between-subjects factors, we calculated eta-squared values, and converted them to f (Cohen, 1988) based on the simple main effects. These analyses yielded the following: for AOE, $\eta^2 = .206$, $f = .1$; for MLOE, $\eta^2 = .207$, $f = .1$. Both these effects sizes can be interpreted as small (Cohen, 1988).

For NDW types, the main effects of the repeated-measure vocabulary scores was significant, $F(4, 28) = 10.06$, $p = .000$, and the between-subjects factor of MLOE, $F(1, 7) = 10.69$, $p = .014$, was significant, but AOE, $F(1, 7) = .927$, $p = .368$, ENGHOME, $F(1, 7) = .294$, $p = .605$, and NVIQ, $F(1, 7) = 2.74$, $p = .142$, were not. There were no significant interactions. Children's NDW types over time are plotted according to their between-subjects MLOE groups in Figure 5. Akin to the results in Figure 4b, the children in the high MLOE group had consistently higher numbers of unique word types in production. To estimate the effect size for AOE, an η^2 value of .03 was calculated, corresponding to $f = .05$, which is below a small effect size (Cohen, 1988).

It is possible that the potential impact of English use in the home and nonverbal IQ was not detectable because they were treated as nominal and dichotomous factors. Perhaps these variables were exerting an effect on the children's vocabulary growth that could be seen if they were treated as continuous. To test this possibility, Pearson correlational analyses were conducted between the average use of English in the home scores, the standard scores from the Columbia Mental Maturity Scale (CMMS), and the outcome measures of PPVT raw scores and NDW types at Round 5. ENGHOME scores did not show a significant relationship with either PPVT raw scores ($r = -.324$, $p = .176$), or NDW types ($r = .125$, $p = .611$). CMMS scores had a significant relationship with NDW types ($r = .509^*$, $p = .026$), but not with PPVT raw scores ($r = .324$, $p = .177$).

DISCUSSION

The purpose of this study was to examine vocabulary development over time in children learning English as a L2 from a variety of L1 backgrounds to understand how they compared to monolingual peers, whether they displayed developmental processes such as the GAP verb phenomenon, and what factors influenced individual differences in their rate of lexical development.

Our first analysis assessed the developmental trend over time for our main measures of expressive and receptive vocabulary. PPVT raw scores showed continuous growth, whereas number of different word types displayed a plateau effect. Perhaps this plateau effect was because of our method of using the same interview questions at each round and the same 100 utterance limit on the sample. Whereas

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

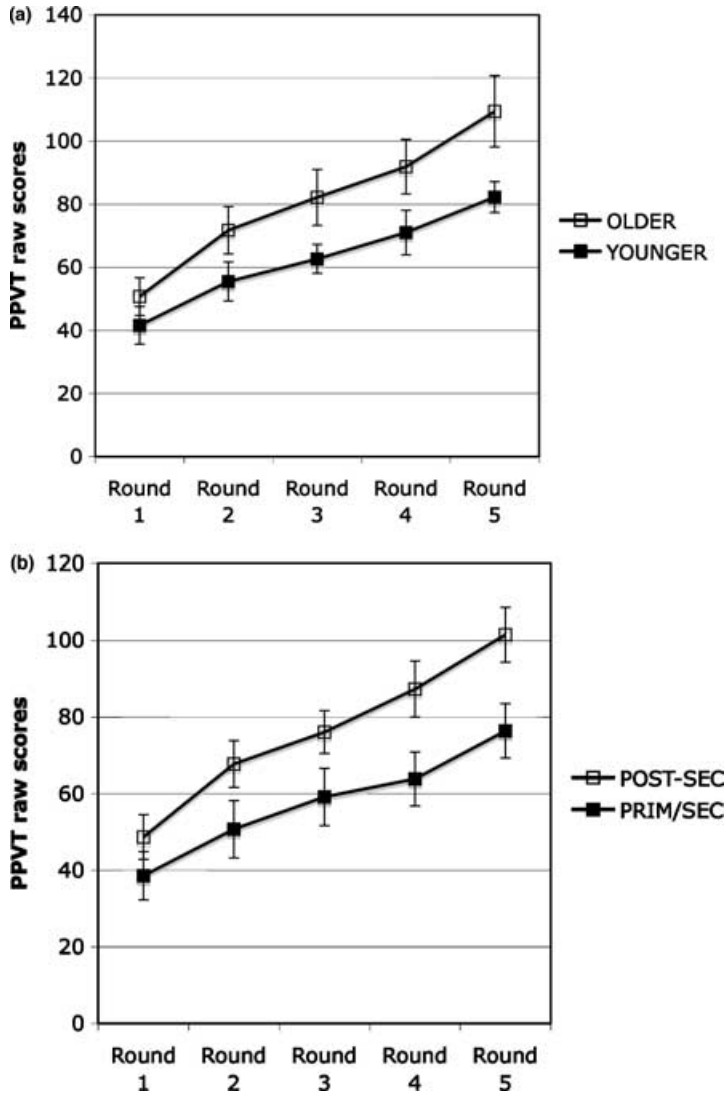


Figure 4. Peabody Picture Vocabulary Test raw scores for five rounds (a) for children with younger and older age of onset of English acquisition and (b) for children of mothers with postsecondary education and with secondary or primary education.

a test such as the PPVT can inherently account for continuous language growth, fixed language sampling may not have this quality. Therefore, we believe that the plateau effect was most likely an artifact of our measure of expressive vocabulary, and not indicative of the general trend in ESL expressive vocabulary development.

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

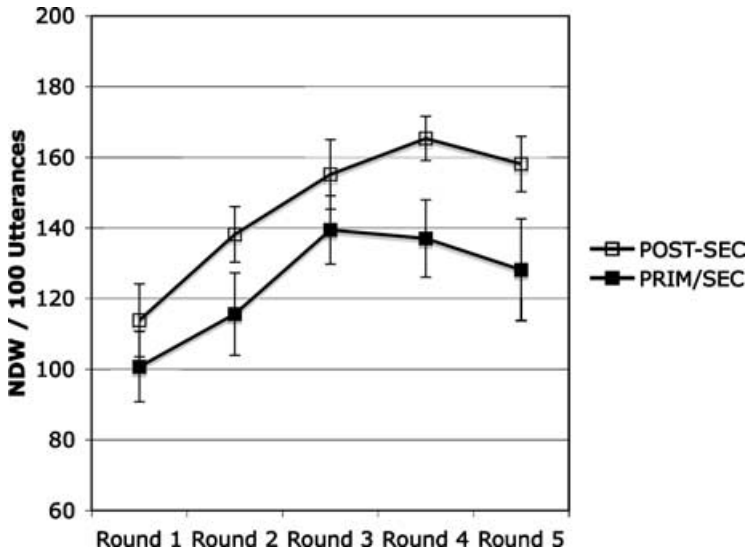


Figure 5. The number of different word types for five rounds for children of mothers with postsecondary education and with secondary or primary education.

Comparing ESL children's and native-speakers' lexical development

We hypothesized that the children in this study would achieve native-speaker norms for vocabulary size faster than the ESL children in the Miami study, based on the contrast in community context between the two groups. The ESL children in Miami all had Spanish as their L1, which is very widely spoken in social and economic spheres in that city and is the language of instruction in some schools (Cobo-Lewis et al. 2002). Therefore, the children in the Miami study would have had extensive opportunities to use their native language with classmates at school, and outside their homes. By contrast, the children in this study resided in Edmonton, Canada, an English majority city with numerous minority languages, none of which is spoken widely enough for the children to have received extensive exposure outside their homes, or to necessarily have had any classmates at school who shared the same L1. Thus, the ESL children in this study were most likely more immersed in English, inside and outside school, than their counterparts in Miami. Considering both standard and age-equivalency scores from the PPVT, the ESL children in this study came close to meeting the norms for age-appropriate receptive vocabulary size in English in just under 3 years' exposure to the language. Recall that the mean standard score at Round 5 was 97, very close to the 100 of the monolingual norming sample. Let us compare the results of our study directly with those of the Miami study at second grade, where those children's exposure to English was similar to that of the children in this study (Cobo-Lewis et al., 2002, pp. 75–76). Among the Spanish at home group in the English-only program, the high SES children's mean standard score was 80, and the low SES

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

children's mean standard score was 65, both lower than the mean score for the ESL children in Edmonton. Even more striking is a comparison with the fifth-grade Miami children, where the high and low SES group means were 93 and 87, respectively, which are still lower than the mean of the children in this study. However, although this comparison with the Miami study indicates that rates of English L2 development vary across community contexts, it does not directly indicate the underlying cause of the variation. Because we did not collect data directly on children's social networks, or use of English outside the home with peers, we can only speculate on the influence of community context differences on children's exposure to English. Furthermore, differences between the results of this study and those of Cobo-Lewis et al. (2002) could be because of other factors, such as differential age of L2 onset in the children in this study, or differences in instructional programs between Miami and Edmonton.

Regarding our measure of expressive vocabulary, lexical diversity in spontaneous speech, we found that at Round 1, the ESL children used fewer different word types in production than their native-speaker age peers, as reported in Watkins et al. (1995). However, it would be unreasonable to expect that after just 9 months' exposure, these ESL children would have caught up. The more crucial comparison with the data from Watkins et al.(1995) is the one based on amount of exposure to English. We found that at Round 5, these ESL children used more different word types in production than younger native speakers who had had the same amount of exposure to English. This second comparison suggests that ESL children outpace younger monolingual peers for expressive vocabulary growth, which is consistent with our analysis of receptive vocabulary, and with findings on verbal academic skills from other studies (Collier, 1987, 1989).

The GAP verb phenomenon in ESL children's lexical development

We examined children's use in context of the verb *do* to determine if it was overextended such that it appeared to be a GAP verb, and if this changed over time. We hypothesized that the GAP phenomenon would appear in the language production of minority L1 children learning English as an L2, but would decline more rapidly than has been reported for L2 learners in French immersion because the minority L1 children had greater and more comprehensive exposure to their L2 inside and outside school. Our results support this hypothesis. GAP uses of *do* by the children were numerous from Rounds 1 to 2, with mean frequencies equal to their appropriate *do* uses, but GAP uses were significantly lower than appropriate uses after that time. In contrast, Harley (1992) found evidence of the GAP phenomenon in fourth grade children who had been in immersion since kindergarten. The presence of GAP *do* uses early on illustrates one way in which these ESL children stretched their meager English lexical resources at that time. The near disappearance of GAP *do* uses by Round 5 ($M = 0.4$) coincided with the children's receptive vocabulary approaching native-speaker age expectations. This coincidence indicates that reliance on GAP verbs could have been related to overall vocabulary size in these L2 learners. Furthermore, note that many of the more specific verbs suggested by the context, given in Table 3, were not low frequency or rare verbs, and it would be surprising given the PPVT scores

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

that the children would not have been familiar with most of these verbs. It is therefore possible that GAP verb usage might have been the result of difficulties in accessing words at the early stages of L2 learning, as well as having been the result of insufficient vocabulary knowledge. In other words, in some instances the more specific word might not have been accessed as quickly as the less specific word, rather than the more specific word not having been known at all.

Sources of individual differences in ESL children's lexical development

The most consistent predictor of children's vocabulary development in English was mother's level of education, which was used as a measure of SES. The ESL children whose mothers had postsecondary education had higher PPVT raw scores and a greater number of different words in production across rounds. Our findings join a number of studies on both monolinguals and bilinguals showing that children's lexical development is enhanced in higher SES families (Arriaga et al., 1998; Cobo-Lewis et al., 2002; Fenson et al., 1994; Hart & Risley, 1992; Hoff, 2003a; Hoff-Ginsberg, 1998; Lambert & Taylor, 1990). Like Cobo-Lewis et al. (2002), SES proved to be a significant variable in English vocabulary development even when this was not the language parents were using most often with the children. In fact, mother's level of education was moderately and negatively correlated with use of English in the home; therefore, the children with highly educated mothers tended to be those getting less English input at home. Findings from this study together with those of Cobo-Lewis et al. (2002) raise interesting issues concerning the reasons underlying why SES makes a difference in children's lexical development in a language being learned mainly outside the home. In particular, these findings suggest that higher order and nonlanguage specific verbal interaction factors associated with SES, such as asking conversation-continuing questions, may be relatively more pivotal than exposure to a larger quantity of individual vocabulary items in the target language for language learning enhancement (but see Hoff, 2003a). Alternatively, the enhancement ESL children of highly educated mothers receive in their L1, both quantitative and qualitative, builds L1 vocabulary mainly, but larger L1 vocabularies might provide a better starting point for learning L2 vocabulary (cf. Cummins, 1991, 2000). A third interpretation could be that mothers with higher education have higher IQs or language learning aptitude and pass these traits onto their children genetically. We do not favor this interpretation for the following reasons. First, the relative contribution of environmental and internal factors has been investigated in the monolingual literature with respect to SES and children's language development, and there is evidence against internal or genetic factors underlying the SES-based differences (see Hoff 2003a). Second, in this study, mother's level of education did not correlate with children's nonverbal IQ scores. Third and finally, the proposal that mothers with higher education are more intelligent than those with less education rests on the assumption that access to postsecondary education is mainly based on talent rather than economic resources and social class. In many of the home countries of the mothers in our study, it is unlikely that access to postsecondary education would have been based on talent alone. In sum, our findings, together with those of Cobo-Lewis et al. (2002), highlight the need for deeper understanding of how mother's level of

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

education influences verbal interaction cross-culturally, and how in turn this may be related to children's language development in both their L1 and their L2.

In addition to mother's level of education, age of onset of English acquisition was also a significant source of individual differences, but only for receptive vocabulary development. Children who began learning English after 5 years of age had higher PPVT raw scores across rounds than children who began learning English before 5 years of age. This finding is similar to results from studies examining age of L2 onset and academic verbal skills, showing a middle-childhood onset advantage (e.g., Collier, 1987, 1989), although note that the mean older age of L2 onset was 5;6 in this study, which is younger than middle childhood. Finally, it is important to point out that because the children in the younger age of L2 onset group were also chronologically younger at each testing interval, the data in this study cannot disentangle whether age of L2 onset or general cognitive maturity was the key factor underlying these age-based effects. Further research on children with differential ages of L2 onset, but with the same chronological age and amount of exposure to English, needs to be conducted to address this issue.

Contrary to Cobo-Lewis et al. (2002), we did not find that English use in the home had an enhancing effect on children's vocabulary development. For example, we did not find that children who were in homes where English was used often showed larger receptive vocabularies and more lexical diversity in production compared with children who were in homes where English was not often used. One possible reason for the difference between our study and Cobo-Lewis et al. (2002) could lie in the parents' fluency in English. Virtually all the parents in this study were recent arrivals to Canada, and only a few were fluent in English. For example at Round 1, mean self-reported English fluency ratings were 2.29 for mothers, 2.50 for fathers on a 1- to 5-point scale, and by Round 5 little had changed as they were 2.50 for mothers and 2.78 for fathers. In contrast, many of the parents who chose to speak English at home in the Miami study were fluent English speakers (Cobo-Lewis et al., 2002). Another possible reason for the limited effect of English in the home could be that, as mentioned above, the families more likely to use English in the home were those where the mother had lower levels of education. It is possible, then, that there were differences in the effectiveness of interactions between parents and children in homes where English was used more often. Thus, we could hypothesize that the additional exposure provided by the use of English in the home was trumped by quality of interactional style in terms of the impact on children's English lexical development.

In our above discussion of the results regarding rate of lexical acquisition, we speculated that the children in this study would receive more opportunity to use English outside the classroom than the Miami cohort because of the different community contexts, and that these contextual differences might result in the children in this study showing a faster rate of development. By contrast, we have just argued that the children in the Miami study most likely had access to superior English input at home, when they received English input in the home at all, and this might explain why English input in the home had a positive effect on children's vocabularies in that study but not in ours. This leaves open the question of why the superior home English versus superior community English exposure would not balance out such that differences between the groups of children in these

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

contexts would be minimized. Although we do not have a satisfactory answer to this question at present, we would like to point out that in developing an answer, the intervening factors of age of L2 onset and SES must be considered, and in addition, data collection on community context factors in Edmonton is needed. For example, Jia (2003) measured the richness of the English environment including factors such as TV, books, and English-speaking friends. She found that these sources of native-speaker input had a positive effect on the English development of Mandarin L1 children in New York, who like the children in this study, had a more minority L1 than the children in the Miami study (see also Jia & Aaronson, 2003).

Like higher use of English in the home, nonverbal IQ as a dichotomous variable did not make a difference in these children's development of receptive and expressive vocabulary in English over time. In contrast, we did find a correlation between nonverbal IQ scores and expressive (but not receptive) vocabulary at Round 5, which could be considered consistent with the findings of Genesee and Hamayan (1980), where general cognitive reasoning abilities were correlated to children's L2 academic verbal skills at the end of kindergarten. However, even though general intelligence is related to language aptitude, it can only be considered an approximate measure of language aptitude (Dörnyei & Skehan, 2003; Sawyer & Ranta, 2001). Therefore, nonverbal IQ may not be specific enough in what it measures to be predictive of children's language learning outcomes as reliably as a true language aptitude test. It is possible that further research with tasks that more precisely measure the working memory and analytic skills thought to comprise language aptitude might find a more consistent relationship with lexical development than we found in this study.

APPLIED IMPLICATIONS AND CONCLUSIONS

We would like to highlight key points from this and other studies that may be relevant to educational and health care professionals. ESL children may take somewhere between 3 and 6 years to be indistinguishable from English native speakers in receptive and expressive vocabulary development. Along the way, ESL children may display patterns of limited verb lexical diversity as they stretch their lexicons to meet the demands of the communicative situation, and in so doing, may overextend semantic meanings of verbs. Individual differences in rates of development are noteworthy, for example, even if the mean PPVT standard score of the ESL children in this study approached 100 after 2.10 years exposure to English, there were three children in the group whose scores did not even reach 85, and there was no reason to suspect any of these three to have had language learning disabilities. Caution is needed when comparing group means to any individual L2 learner's performance. Slower than average L2 learners might be misdiagnosed as having language or learning disabilities if educators and speech–language pathologists are not aware of how broad the range of typical L2 development can be (see also August et al., 2005; Genesee, Paradis, & Crago, 2004; Paradis, 2005).

Certain factors have been found to be associated with individual differences in rates of L2 development in this and other studies. For example, children from higher SES backgrounds who were older than preschoolers appeared to have some

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

advantage. Higher general intelligence exerted a limited effect as well. These factors could be considered in setting appropriate expectations for an ESL child's English development. However, in the present study, the effect sizes for these sources of individual differences were small. Although small effect sizes are the norm in behavioral science research (Cohen, 1988), small effect sizes mean that the ability to generalize from these group findings to individual children is limited. This is particularly important for educational policy decisions and giving advice about when it is best for minority language children to begin learning English (cf. Wong Fillmore, 1991). Further, we should reiterate that because chronological age and age of L2 onset were not separated in this study, this design makes it impossible to claim from these data that waiting until school age to introduce English is the best practice.

The nonenhancing effect of English use in the home in this study is relevant to the debate over whether English in the home should be encouraged for minority language children, and what influence this might have on maintenance of the L1. Educators and healthcare professionals may encourage use of English in the home to give the child maximal exposure and practice with the majority language of the society (Genesee et al., 2004). In contrast, researchers have argued that L1 maintenance is directly related to language choice in the home, and L1 maintenance is beneficial to minority children's cognitive, educational and social–emotional development (Anderson, 2004; Cummins, 1991, 2000; Hakuta & D'Andrea, 1992; Wong Fillmore, 1991). The notion that speaking English in the home boosts English lexical development because it increases children's English input was not supported in this study; however, contrastive findings come from other studies, for example, Cobo-Lewis et al. (2002). Further research is needed to more thoroughly understand the effects of use of English versus the native language in the home, and whether the effects of English use in the home are mitigated by other factors, such as SES, parental fluency in English, or community context. It is our opinion that research can only adequately inform best practices with respect to this issue after a more thorough understanding has been achieved.

ACKNOWLEDGMENTS

We thank the children and their families for their dedication to this longitudinal study and the following organizations who helped us recruit participants: ABC Head Start, NorQuest College English Language Training, ASSIST Community Services Centre, The Mennonite Centre for Newcomers, and The Multicultural Healthbrokers. We also acknowledge the student assistants who collected the data: Lisa Brown, Julie Coutu, Tamara Duncan, Lindsay Griener, Laura Marcon, and Cinnamon Suyal. This research was supported by the Social Sciences and Humanities Research Council of Canada (Standard Research Grant 410-2002-0093) and the Alberta Heritage Foundation for Medical Research (Establishment Grant 200100341), for which we are grateful.

NOTES

1. Two of the children received some instruction in Mandarin at school, but we have no reason to suspect it significantly affected their English development because these

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

children did not perform similarly to each other or systematically differently from the children in English-only programs, and furthermore, one spoke Cantonese at home.

2. In Paradis (2005), the reported use of English in the home at Round 1 was parents' use only. When parents indicated how much they spoke English to their children, this was typically lower than English use in the home by older siblings. For this reason, reported use of English in the home at Round 1 was lower in Paradis (2005) than in the present study.

REFERENCES

Anderson, R. (2004). First language loss in Spanish-speaking children: Patterns of loss and implications for clinical practice. In B. Goldstein (Ed.), *Bilingual language development and disorders in Spanish–English speakers* (pp. 187–212). Baltimore, MD: Paul Brookes.

Arriaga, R. J., Fenson, L., Cronan, T., & Pethick, S. J. (1998). Scores on the MacArthur Communicative Development Inventory of children from low- and middle-income families. *Applied Psycholinguistics, 19*, 209–223.

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English Language Learners. *Learning Disabilities Research & Practice, 20*, 50–57.

Bialystok, E., & Miller, B. (1999). The problem of age in second language acquisition: Influences from language, structure and task. *Bilingualism: Language and Cognition, 2*, 127–145.

Burgemeister, B., Hollander Blum, L., & Lorge, I. (1972). *The Columbia Mental Maturity Scale*. New York: Psychological Corporation.

Carroll, J., & Sapon, S. (1959). *The Modern Languages Aptitude Test*. San Antonio, TX: Psychological Corporation.

Cobo-Lewis, A. B., Pearson, B. Z., Eiler, R. E., & Umbel, V. C. (2002). Effects of bilingualism and bilingual educations on oral and written English skills: A multifactor study of standardized test outcomes. In D. K. Oller, & R. E. Eilers, (Eds.), *Language and literacy in bilingual children* (pp. 43–63). Clevedon: Multilingual Matters.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Collier, V. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly, 21*, 617–641.

Collier, V. (1989). How long? A synthesis of research on academic achievement in a second language. *TESOL Quarterly, 23*, 509–531.

Conti-Ramsden, G., & Jones, M. (1997). Verb use in specific language impairment. *Journal of Speech, Language and Hearing Research, 40*, 1298–1313.

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In Bialystok E. (Ed.), *Language processing in bilingual children* (pp. 70–89). Cambridge: Cambridge University Press.

Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon: Multilingual Matters.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford: Blackwell.

Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59* (Serial No. 242).

Flege, J. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 101–132). Mahwah, NJ: Erlbaum.

Genesee, F., & Hamayan, E. (1980). Individual differences in second language learning. *Applied Psycholinguistics, 1*, 95–110.

Genesee, F., Paradis, J., & Crago, M. (2004). *Dual language development and disorders: A handbook on bilingualism and second language learning*. Baltimore, MD: Paul Brookes.

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development, 1*, 23–64.
Hakuta, K., & D'Andrea, (1992). Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics, 13*, 72–99.
Harley, B. (1992). Patterns of second language development in French immersion. *French Language Studies, 2*, 159–183.
Harley, B., & King, M. (1989). Verb lexis in the written compositions of young L2 learners. *Studies in Second Language Acquisition, 11*, 415–439.
Hart, B., & Risley, T. (1992). American parenting of language-learning children: Persisting differences in family–child interactions observed in natural home environments. *Developmental Psychology, 28*, 1096–1105.
Hoff, E. (2003a). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*, 1368–1378.
Hoff, E. (2003b). Causes and consequences of SES-related differences in parent-to-child speech. In M. Bornstein & R. Bradley (Eds.), *Socioeconomic status, parenting, and child development* (pp. 147–160). Mahwah, NJ: Erlbaum.
Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In M. Bornstein (Ed.), *Handbook of parenting: Vol 2. Biology and ecology of parenting* (pp. 231–251). Mahwah, NJ: Erlbaum.
Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics, 19*, 603–629.
Jia, G. (2003). The acquisition of the English plural morpheme by native Mandarin Chinese-speaking children. *Journal of Speech, Language, and Hearing Research, 46*, 1297–1311.
Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics, 24*, 131–161.
Lambert, W., & Taylor, E. (1990). *Coping with cultural racial diversity in urban America*. New York: Praeger.
MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.
Paradis, J. (2005). Grammatical morphology in children learning English as a second language: Implications of similarities with specific language impairment. *Language, Speech and Hearing Services in Schools, 36*, 172–187.
Peña, E., Bedore, L. M., & Rappazzo, C. (2003). Comparison of Spanish, English, and bilingual children's performance across semantic tasks. *Language, Speech & Hearing Services in the Schools, 34*, 5–16.
Rice, M., & Bode, J. (1993). GAPS in the verb lexicons of children with specific language impairment. *First Language, 13*, 113–131.
Rice, M., & Wexler, K. (2001). *Test of early grammatical impairment: Examiner's manual*. New York: Psychological Corporation.
Saunders, W., & O'Brien, G. (in press). Oral language. In F. Genesse, K. Lindholm-Leary, W. Saunders, & D. Christian (Eds.), *Educating English language learners: A synthesis of empirical evidence* (pp. 24–97). New York: Cambridge University Press.
Sawyer, M., & Ranta, L. (2002). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319–353). Cambridge: Cambridge University Press.
Thordardottir, E., & Ellis Weismer, S. (2001). High-frequency verbs and verb diversity in the spontaneous speech of school-age children with specific language impairment. *International Journal of Language & Communication Disorders, 36*, 221–244.
Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Umbel, V. M., Pearson, B. Z., Fernández, M. C., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development, 63*, 1012–1020.
Verhallen, M., & Schoonen, R. (1993). Lexical knowledge of monolingual and bilingual children. *Applied Linguistics, 14*, 344–363.
Vermeer, A. (2001). Beadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics, 22*, 217–234.

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press

Watkins, R. V., Kelly, D. J., Harbers, H. M., & Hollis, W. (1995). Measuring children's lexical diversity: Differentiating typical and impaired language learners. *Journal of Speech and Hearing Research, 38*, 1349–1355.

Weber-Fox, C., & Neville, H. (2001). Sensitive periods differentiate processing of open- and closed-class words: An ERP study of bilinguals. *Journal of Speech, Language and Hearing Research, 44*, 1338–1353.

Windsor, J., & Kohnert, K. (2004). The search for common ground: Part I. Lexical performance by linguistically diverse learners. *Journal of Speech, Language and Hearing Research, 47*, 877–890.

Wong Fillmore, L. (1991). When learning a second language means losing the first. *Early Childhood Research Quarterly, 6*, 323–346.

https://doi.org/10.1017/S014271640808003X Published online by Cambridge University Press