NIH R21--2006 Tom Roeper, P.I. Barbara Zurer Pearson, Research Associate
(Revision of 2005 proposal)
Abridged 9/22/2008

Understanding quantifiers in language and mathematics: Age and dialect differences.
A. Specific Aims

Research in semantic development has established that children and second language learners often understand quantifier words, like *all*, *some*, and *every*, in ways that differ systematically from native-speaker adults. Much of the research has focused on specific semantic properties of quantifiers, (e.g. how they take scope or can be displaced), which appear to give rise to children's non-target interpretations. For example, when asked to choose a picture where "some boxes are black," many children rejected one with some all-black boxes in favor of a picture with partly black boxes, as if the sentence had an additional *some*: "some of the boxes have some black"(Roeper & Matthei, 1978). Copying a quantifier to a different part of the sentence looks like a strange misconception to English speakers, but represents a common syntactic process in many languages of the world. We hypothesize that such interpretation strategies, which persist into middle childhood (Seymour et al. 2002; Roeper et al. 2006), create ambiguities for children in the wording of math problems that adults do not experience and thus do not anticipate. They may impair children's ability to do math problems for linguistic, not mathematical reasons. Spanish and African-American English (AAE) have quantifier systems that are different from mainstream American English (MAE). We therefore predict that learners of MAE who come from those language backgrounds will perceive a different range of ambiguities. From the other perspective, some of children's problems with quantifiers may depend on achievement of specific concepts in mathematics: most, for example, includes an implicit comparison which may rely on the child's ability to estimate inequalities.

In order to better understand the relation between semantic interpretations of specific quantifiers and mathematics performance, a fine-grained picture of the steps and approximate timetable of acquisition for a range of semantic phenomena is required. The acquisition of the semantic properties of a larger range of quantifiers than has been studied thus far must be undertaken, and the quantifiers should be studied in the context of children's progress in mathematics. Both acquisition paths (mathematics and language) should be explored with a more diverse sample of children.

We propose to study four semantic properties (exhaustivity, distributivity, concord, and displacement/ or spreading) with respect to the quantifiers, *all, every, each, some*, and *most.* The subject groups will be learners at six ages between 4 and 14 from three dialects of English: MAE, AAE, and Spanish-influenced English. Each dialect group will have a control group of adults. Our language probes will elicit comprehension and production data using two question types: 1) sentences describing sets of geometric shapes (like Inhelder & Piaget, 1964) and 2) quantified expressions integrated into story problems anchored in the world of real objects and people. The math probes will be standard diagnostic assessments, such as those administered at the beginning of each school year, which give both a total score and a profile of which concepts and computations a child has mastered and which ones will be the targets of instruction. "Think-aloud" data will also be collected and analyzed in conjunction with the quantitative data.

The deliverables produced by the study will be

1) developmental milestones in each dialect for each of the five quantifiers (with respect to the semantic properties discussed below);
2) statistical analyses (especially correlation and regression) of the relation between the developmental milestones for the quantifiers and the mathematics probes; and
3) for at least a few of the semantic properties, a sketch of a causal relationship between quantifier understanding and specific mathematics achievements, for example one-to-one counting and distributivity, or the comparative most and the estimation of inequalities.

Such a study will have implications for math education for all children, but most especially for children of color. (Note that both ELL and AAE speaking children are here called "learners of MAE" as they are both acquiring MAE after acquisition of a first language or dialect. They are to be distinguished from those learning MAE as a first dialect, whom for the purposes of comparison we will call "native MAE-speakers.")

The work is innovative in that it applies tools of analysis from language acquisition studies and formal semantics to issues in mathematics education and vice versa. It attempts the difficult translation of terms and concepts that are grounded in one discipline into terms and techniques which (one hopes) will be useful in another discipline. This very focused topic at the intersection of language and math holds the promise of contributions in both arenas. Furthermore, it focuses on two populations which are difficult to study (Craig, 1996; Oller & Eilers, 2002b), but where the need for greater knowledge is critical.

B. Background and Significance
The relationship of language and math:

Mathematics and language are generally considered distinct skills (Varley, Klessinger, Romanowski, & Siegal, 2005) to be tested and reported separately. But there is a growing awareness that language development may have profound effects on how children's math concepts develop (Lakoff & Nunez, 2000; Mestre, 1988; Riley & Greeno, 1988; Gelman & Gallistiel, 2004). Word problems, in particular, engage both sets of abilities.

Consider this simple question:
        Show a picture of two boys, and ask:
        "Do they have two hands or four hands?"

If I count the hands, I get four. If I understand the question to include an implicit *each*, as most adults probably would, ("Do they each have two hands or four hands?") then the answer is two. If the child says "three," it is a mathematics problem, but whether she says "two" or "four" is a question of language.

In recent years, children's comprehension of sentences containing quantified expressions has received considerable attention from investigators in the fields of psycholinguistics and developmental semantics (Takahashi, 1991; Philip, 1995, 1996, 2004; Brooks & Braine 1996; Crain, Thornton, Boster, Conway, Lillo-Martin & Woodams, 1996; Crain, 2000; Krämer, 2000; Drozd, 2001; Lidz & Musolino, 2002; Guerts, 2002 among many others). The observation that children often differ from adults in the way they interpret sentences containing quantified expressions is well-attested in many different contexts and in several different languages (Philip, 1995, 1996). While the factors underlying the various phenomena are still being debated, the validity of this observation is beyond dispute (Roeper, Strauss, & Pearson, 2006). There are considerable differences between children's and adults' interpretations, and they have been shown to persist a surprisingly long time.

We have only recently realized the potential application of this work to the field of math. While children are working out the grammar of these sentences throughout the elementary school years, there may be ambiguities for them in situations where adults do not experience ambiguity, or conversely, children may not compute alternate interpretations that are natural and necessary for adults. Most people not acquainted with the psycholinguistics literature express surprise at children's (mis)interpretations of quantified sentences. Even math teachers who have written articles and books on the relation between language and math (e.g. Spanos, Rhodes, Dale, & Crandall, 1988; Hirigoyen, 1997; Khisty, 1997; Murray, 2004) appear not to have focused on quantifiers and give no sign of being aware of the ambiguities they create in the wording of math problems like those on the high-stakes exams.

Historical background

Children's difficulties with the quantifiers *all* and *some* have been well known since the work of Piaget (Inhelder & Piaget, 1964). Piaget showed children ages 5 to 9 different arrays of red and blue circles and squares and asked questions like, "Are all the circles blue?" "Are all the blue ones circles?" His astounding results showed that through middle childhood, children's interpretations of quantifiers were systematically different from adult interpretations. Shown an array of three blue circles, one blue square, and two red squares and asked: "Are all the circles blue?" a large percentage of the children through age 8 said "no." They would point to the blue square and say "not this one" or "not this square," as if the question were: "Are all the circles all of the blue things?

Similar results have been found with *every*. When asked: "Is every child riding a bike?" a majority of children through age 8 said, "no, not this bike," when there was an extra bike, but no extra child (Philip, 1995; Drozd, 2001; Seymour et al. 2002). Because the *all* or *every* appears to spread from where it is in the surface structure to another part of the sentence, it is often called "quantifier spreading" (Philip, 1995; Roeper, Strauss, & Pearson, 2006). A second type of spreading ("type-2") has also been observed, where

every seems to mean "always": Is it always the case that a child is riding a bike? When children are shown a picture of 3 children riding bikes and a rabbit riding a scooter, and are asked "Is every child riding a bike?" in type-2 spreading, the child would point to the rabbit and answer, "not the rabbit" or "not the scooter" even though neither rabbits nor scooters were mentioned or asked for in the question (Guerts, 2002; Roeper, Strauss, & Pearson, 2006).

Linguistic dimensions of the differences

The difficulties children have in assigning adult-like interpretations to words like *some* and *all* have often been attributed to conceptual immaturity (Inhelder & Piaget, 1964) or have sometimes been considered an artifact of the probes (Crain et al., 1996). However, the demonstration by dellaCarpini (2003) that adult second language learners of English exhibit many of the same interpretation strategies with respect to quantifier scope that children do suggests it is not solely a factor of child cognition nor language-independent child pragmatics. Rather learners must determine from the available linguistic input how the semantic possibilities presented by quantifiers are encoded in the target language. The spreading phenomenon, which looks like a misconception to adults, appears in the grammar of many different languages in various ways. From the point of view of world languages, it is a relatively common syntactic process. As we shall see, it is also found in the MAE quantifier system, just not with *every*.

Other quantifier displacements in English.

A learner's spreading hypothesis might also be reinforced by the fact that, like adverbs, some other quantifiers can "legally" occupy several different positions in the sentence without changing the essential meaning. Consider:

1a. "The boys have two hats <u>each</u>" and
1b. "<u>Each</u> of the boys has two hats."
2a. "<u>Both</u> boys left early" and
2b. "The boys <u>both</u> left early."
3a. "<u>All</u> the children were singing in the concert" and
3b."The children were <u>all</u> singing in the concert."

The (a) and (b) versions of the sentences both describe the same situations. One might even say,

1c. "<u>Each</u> of the boys has two hats <u>each</u>."

This might appear redundant, but the meaning does not change. By contrast,

4a. "A cat climbed <u>every</u> tree"
4b. "<u>Every</u> cat climbed a tree," and
4c. "<u>Every</u> cat climbed <u>every</u> tree."

all describe different relationships of cats to trees.

Native and non-native-MAE learners alike must recognize that *every* does not move as freely as *each* and *both* and *all*. It appears that most MAE-learning children (and L2 learners) eventually learn the proper constraints on *every*, but it takes a surprisingly long time. Children use *every* early in expressions like *everyone* and *every day,* so without careful testing, we are not aware of the limits on their understanding of it in different contexts. In a study of over 1,000 typically developing children (Seymour et al. 2002), we were very surprised to observe that the percentage of spreaders <u>rose</u> between 5 and 8 years of age, and among 12-year-olds, the oldest subjects in the study, over 25% were still spreading. It is quite plausible for errors such as these to make for math difficulties that teachers may not be alert to. Even when we are aware of them, we cannot completely eliminate them, but at least we can take them into account when testing or creating tasks for children.

The collective/distributive difference:

Another semantic property with the potential to trip children up in math while they are still working out the grammar is the distributive versus collective distinction. Compare:

a) I lifted <u>all</u> the rocks

      b) I lifted <u>each</u> rock.
      c) I lifted <u>every</u> rock.

In (a) a person can either lift the whole group of rocks, or he can lift them one by one. In (b) it must be the case that he lifted the rocks one by one. Initially, *every* (in c) seems like *each*, but it could also describe a situation where the rocks were all lifted at once <u>or</u> one by one (Tunstall, 1998). If we say, "All the flowers were in a vase," we probably mean in one vase, but for "each flower was in a vase" there are probably as many vases as flowers. For *every,* we predict that adults would choose each scenario (one vase or many) about 50% of the time.

      If we say "There are three boys. Every boy has three hats. How many hats are there?" the answer is most likely 9. But a child who thinks *all = every* might understand the situation as "<u>all</u> the boys (together) have three hats." In that case, "three" might not be exactly wrong. Some wordings make this construal very natural. If children first think *every* here equals *all* and they read it collectively, then they will give the wrong answer to the math question although it is a correct interpretation in their grammar.

      On the other hand, children sometimes apply distributivity when it is not warranted. Consider:

      "There is a horse that every child is on."

de Villiers & Roeper (1993) showed children two pictures for this sentence: in one there was one horse with three children on it; in the other, three horses each with a boy on it. Children aged 6 chose the pictures equally, whereas for adults, the first reading, which does not distribute boys to horses, is preferred. If the sentence were "Every child is on a horse," it would be fine to understand it as three horses, one for each boy. Here it is a grammatical barrier, the relative clause, that does not allow the *every* to apply to *horse.* Thus we have another example of grammar, not conceptual development *per se*, restricting the child's numerical understanding of the situation described by the quantified sentence.

Other relevant semantic properties of quantifiers: exhaustivity, other displacements
      Spreading and distributivity are just two of the semantic properties that can give rise to misinterpretations of quantified sentences by MAE-learners. Exhaustivity—the requirement to enumerate all the members of a mentioned set—has also been shown to develop slowly and at different rates in children learning different languages (Roeper, Pearson, Penner & Schulz, 2002; Roeper et al. 2005). Similarly, displacement, especially as discussed by Lidz and Musolino (2002), is another potential source of misunderstanding of quantified sentences, in their case, quantified negative sentences: (eg. "Every student ca<u>n't</u> afford a car" = "Not every student can afford a car").

      It is important to note that quantifiers are also present in some words that are not generally considered quantifiers: *wh*-words, (*who, what, where, when, which*). These words are actually composites: *wh + then = when, wh + here = where.* In the same way, *wh* + the quantifier *each* equals *which* (Roeper, in press). In addition, all *wh*-words in English include an implicit *every*, in that they require set responses, which in most contexts must be exhaustive. (If one asks, "Who was in the car the night of the murder?" the answer must give all the occupants, not just a few.) By contrast, the Japanese language has two separate *wh*-words, one for "who-someone" when the expected answer is an individual, and one for "who-everyone" when it is a set (Nishigauchi, 1999). Learners do not know from the outset whether they are in a Japanese-type language or an English-type language in this respect. Children learning English (and German) do not automatically assume *who* requires an answer of more than one individual. As one 6-year-old told us when asked "who was wearing a hat" (from an array with several boys wearing hats and others without), "I don't know which one to tell you" (Roeper, Pearson, Penner & Schulz, 2002).

Dialect differences:
      Quantifiers are found in many places in a language, and languages differ in how quantifiers interact with other parts of the grammar (Bach, Jelinek, Kratzer, & Partee, 1995). Given the variety of ways that the languages of the world treat quantifiers, we hypothesize that careful examination will reveal that dialects <u>within</u> languages differ along similar dimensions. With respect to spreading, for example, AAE and Spanish have significant grammatical differences elsewhere in the grammar which may affect the acceptability of quantifier displacements like spreading, even among fluent adult speakers of AAE and Spanish-influenced English. One important feature of both AAE and Spanish, but not MAE, is "negative concord." In a

negation system that uses concord (or "agreement"), multiple markers of negation reflect a single instance of negation in different parts of the sentence. The Spanish sentence:

"<u>No</u> tiene <u>ningun</u> amigo" [*he doesn<u>'t</u> have <u>none</u> friend*]

would be translated into MAE as "He has no friend(s)" <u>or</u> "he doesn't have (any) friends" with the negative marker in one place or the other (in the noun phrase or the verb phrase) but not both. It would translate more directly into the AAE sentence with two negation markers, "He don't have no friends." The "same" sentence in MAE represents it as two negations and means he has <u>some</u> friends, as can be made clear by adding a conjunct like "He doesn't have no friends, just very few.") AAE, like Spanish (and many other languages) uses a system of multiple negation. Since it is a common feature in the world's languages, multiple negation would be a reasonable expectation for learners to operate with, and so it will not be surprising if learners of a non-concord dialect like MAE use it until they have enough evidence from the input to reject it. In the meantime, for perfectly sensible reasons, those learners would misinterpret math sentences in word problems.

It remains an unexplored empirical question whether MAE learners from linguistic backgrounds with a system of negative concord will accept spreading for *every* or perhaps attend to other cues that can affect how we know what quantities a sentence is referring to.

<u>Literature on quantifiers in AAE:</u> Orr (1987) raised the problem of dialect differences in math sentences in her book *Twice As Less*, which is based on many examples gleaned from students' "long-hand" descriptions of their problem-solving strategies for algebra word problems. The book focuses mainly on prepositions and expressions of comparison, although there is a short section where the author speculates on how the students understand the quantifiers *all, some,* and *none.* Orr argued that her AAE-speaking students encountered difficulty when they used structures that attempted to merge AAE and MAE. Although she appears to have consulted the work of linguists Labov (1972) and Wolfram & Fasold (1974), there appears to be little linguistic basis for her arguments. For example, she calls double negatives in AAE a mismatched "composite" of two MAE structures (p. 148). Thus it is difficult to integrate her arguments into the current linguistic debate on quantifiers.

Still, the thesis of the book, that the language structures through which children interpret mathematical sentences can create obstacles to math achievement that are not based on math ability, but on language, is in some way the inspiration for this project. However, one needs to pursue the study from a more solid linguistic foundation, as represented by the members of our advisory board (detailed below).

We report below (in Preliminary Studies) on new evidence about AAE-speaking children's interpretations of *every* from Seymour, Roeper, de Villiers, de Villiers, & Pearson (2002). We have published on the MAE data from the same study (Strauss, Roeper, Pearson, de Villiers, & Seymour, 2003; Roeper, Strauss, & Pearson, 2006), but the AAE data have not been published. Nor are there any other studies that our AAE advisor Seymour nor AAE expert Lisa Green could point us to. In brief, the literature on quantifiers in AAE (other than *none*) is very sparse.

<u>Literature on quantifiers in Spanish-influenced English:</u> Mestre (Mestre, Gerace, & Lochhead,1982; Mestre, Thibodeau, & Gerace, 1988) was doing much the same for Spanish as Orr did for AAE, investigating the effect of contrastive features like double negatives on the performance of Hispanic and non-Hispanic college students on word problems with more than one negative. In Mestre's results, language background made a difference in speed and accuracy both at the time of testing and in retention after training.

We know of no studies of Spanish-background learners of MAE which focus on how they recognize distributive and collective readings, when they show awareness of exhaustivity, or how they interpret *every.* One obvious reason for this is that Spanish has words for *all* and *each*, but no word for *every.* Spanish speakers, when talking in English, sometimes choose *each* for "everyday" ('each day' = *cada dia*) and less often *all* ('all the days' = *todos los dias).* For "everytime" they typically say 'all times' = *todas veces.* An analogous situation, for which there is quasi-experimental evidence, may be found in Hispanic-American children's learning of *many,* another English quantifier that does not have a simple counterpart in Spanish. Gathercole's studies of Spanish monolinguals and bilinguals acquiring *many* (Gathercole 1997; 2002) suggest the hypothesis that dialect of origin <u>will</u> make a difference in children's interpretations of quantifiers and their knowledge of the semantic behavior of the quantifiers in the sentences where they occur. Her

work is one of the few studies in the literature of ELL quantifier learning at the level of detail of these subtle semantic and syntactic (and perhaps pragmatic) properties. Much more exploration needs to be done.

Literature on the relation between language and math. Reports from the Early Childhood Longitudinal Study (and other national statistics) (Denton & West, 2002) show generally that children with resources for early success in reading are also the ones who have resources for success in mathematics. Similarly, Klibanoff et al. (2006) in the most recent issue of Developmental Psychology report that their subjects' mathematical knowledge was significantly related to the amount of teachers' math-related talk. The relationships noted are very general. It will be important to explore the role of quantifiers more specifically. Our German colleague Zvi Penner (2006), who has collaborated on many of our studies of quantifiers, has done an exploratory study linking children's accuracy with prepositions and questions to their calculation abilities. There appear to be many avenues to pursue in this arena.

The need for collaborative research:
   The relation between semantic properties of quantifiers and mathematics achievement is a fertile field for experimental investigation. Our work as linguists will benefit from collaboration with mathematics educators who can help insure that the work retains relevance to actual problems children face in their classes. The project therefore calls for close collaboration between investigators in mathematics and investigators in linguistics, both with close ties to minority communities.
   The **research questions (RQ)** to be addressed by the math-language team are as follows:

1. What are the steps and the approximate timetable in native-MAE learners' acquisition of the following four semantic properties --exhaustivity, distributivity, concord/scope, and displacement—as expressed by the quantifiers *all, every, each, some,* and *most.*
2. Are the developmental stages in understanding these quantifiers reflected in children's math skill? Is understanding of the semantic properties of quantifiers related to success in math generally, and also with specific mathematics concepts?
3. Are the steps and approximate timetable for the acquisition of the five quantifiers and their semantic properties the same or different for speakers of AAE? for speakers of Spanish-influenced English?
4. Are the associations between the stages of quantifier learning and mathematics achievement (RQ 1 and 2) different for speakers of native MAE versus AAE versus Spanish-influenced English)?

C. Preliminary Studies

   This project is a natural outgrowth of previous work by the PIs as members of the UMass Working Groups on AAE. It also builds on the co-PI's experience as coordinator of the University of Miami Bilingualism Study Group (BSG), where she worked extensively with Hispanic-background English Language Learners. There are several preliminary studies by the PIs and the Advisory Board members which show that they have successfully engaged in experimental work of this type and which has led them to expect there to be substantive dialect differences.
   The PI has been a leader of an ongoing collaboration of linguists, psychologists, and scholars in Communication Disorders which has fostered scholarly inquiry on AAE since the group came together around 1990 (as advisors on T. Wyatt's dissertation). At least ten dissertations on AAE, several smaller articles, and many presentations at professional conferences have come out of the Working Groups. Most recently they published two test whose goal is unbiased language assessment of children who speak AAE: the *Diagnostic Evaluation of Language Variation (DELV)-Screening Test* and *DELV-Norm Referenced (NR)* (Seymour, Roeper & de Villiers, 2003, 2005). A comprehensive description of the background and development of the *DELV* is published as the February 2004 *Seminars in Speech and Language* guest edited by Seymour & Pearson, with significant contributions from Roeper and de Villiers.

Studies relating to quantifier acquisition in the context of different languages and dialects.
   There are four areas of study by the research group in which Roeper and Pearson are participants (with consultants Seymour, de Villiers, and Perez) which are relevant to this proposal:

1) work showing the developmental stages in the protracted course of quantifier acquisition;
2) work showing a relationship between other linguistic phenomena and quantifier learning;
3) work indicating the existence of dialectal influences on the interpretation of quantifiers; and

4) work showing the connections of the PI and advisors to minority communities.

Studies on quantifier acquisition. The pilot work for the standardized language test *Diagnostic Evaluation of Language Variation (DELV,* Seymour, Roeper & de Villiers, 2003, 2005*)* developed in our lab in conjunction with the publisher Harcourt Assessments, Inc., tested over 20 linguistic constructs, including quantifiers in a nationwide sample of 1500 children. The sample was stratified by age, from 4 to 12 years, so it produced valuable developmental information about a wealth of linguistic elements. A surprising result emerged from the quantifier items included on the pilot. The spreading phenomena known to us among younger children was found with children up to age 12. Further "type-2" spreading was shown to be much more widespread than had previously been observed (Philip, 1995; Guerts, 2002). Analyses of the data from typically developing children, summarized in the "Acquisition Path of Determiner Quantifiers" (Roeper, Strauss, & Pearson, 2006) confirmed that more study of quantifier learning is warranted up through middle school.

Roeper and Pearson have also investigated other aspects of quantifiers. For example, exhaustivity with "*who*" is being explored in a set of experiments carried out jointly with colleagues in Germany. This work (reported in Roeper, Pearson, Penner & Schulz, 2002; Roeper et al. 2005) also uncovers a surprisingly long time during which children's interpretations of the covert quantifier in *wh*-words are different from adults'. Here we found that most children appear to have to learn to answer *wh*-questions with a set-answer, rather than just an individual (unless it is a set of one). (Penner and Schulz are also pursuing studies in German similar to the current proposal.)

Studies on the relation of quantifiers to other aspects of language. Related studies, using the data from the *DELV* pilot (Seymour et al. 2002) showed a strong correlation between children's understanding of exhaustivity in *wh*-words and their success with the word "*every*." Strauss, Roeper, de Villiers, Pearson, & Seymour, (2003) found that among all ages and dialect groups, the children who correctly answered the *wh*-questions exhaustively were significantly more likely to understand the quantifier *every* in an adult way and vice versa. This work, too, suggests strongly that there are quantitative aspects of language use in unexpected parts of the grammar and that it will be fruitful to explore that link further.

Studies on dialect differences in the comprehension of quantifiers. The *DELV* project of the research team also uncovered dialect difference in the comprehension of quantifiers. Items involving the quantifier *every* started out on the diagnostic portion of the test in the pilot studies, and indeed many aspect of children's understanding of *every* is the same in MAE and AAE. The two dialect groups showed equivalent knowledge (and a similar growth curve) of sentence boundary constraints on *every* and in understanding several of its uses where spreading was not an issue. Those items survived the piloting and appear on the *DELV-Norm Referenced Test (Seymour et al., 2005).* But the pilot studies (Seymour et al., 2002; Roeper, Strauss, & Pearson, 2006) showed that MAE-learning children were more likely to show classic spreading (as Piaget found), whereas the AAE-speaking children were more likely to do Type-2 or "always"-type spreading (unpublished).

Work by consultant Johnson on cross-linguistic differences in verb phrase ellipsis and by Perez on ellipsis understanding in Spanish children further suggests avenues for dialect differences in the problems encountered in mathematics materials. An important experiment by Perez et al. (1999) on "generics (that is, uses of bare plurals versus the determiner, "lions" versus "the lion" [are/is the king of the jungle] in Spanish and English will also help us anticipate added ambiguities for Spanish-English bilingual children.

Studies with linguistically and culturally diverse groups. Finally, much previous work by the PI and his consultants shows that they are familiar with linguistically and culturally diverse populations. In the course of test development for the *DELV*, the research team (which included Pearson as Project Manager for 6 of the 12 years of the study) had direct contact with over 1,000 child AAE speakers in the Springfield MA-Hartford CT region. The *DELV* team had indirect contact with over 2,000 more AA children who participated in the standardization research conducted by Harcourt Assessments (formerly TPC) for whom the Working Groups did all the linguistic analysis of the voluminous data. The current proposal also relates to previous studies by Pearson with Spanish/English bilingual children during the 10 years she was the Coordinator of the University of Miami Bilingualism Study Group (Oller & Eilers, 2002a; Pearson, Fernandez & Oller, 1993; Pearson, 1993). In addition, five of the consultants (Elliott, Hill, Perez, Seymour, & Wallace) are persons of color and have both professional and personal ties to the African American and Hispanic communities.

Preliminary studies surveying mathematics word problems. [THIS SECTION IS NEW.]

The members of the Advisory Board for the current proposal have extensive familiarity with mathematics education from a national perspective and a specific focus on word problems. One consultant, M. Pearson, is Associate Director of the national organization, the Mathematics Association of America (MAA), and P. Elliott is a leader in the National Council of Teachers of Math (NCTM). An examination by Roeper (in press) of the practice problems for the Massachusetts 4th grade state-wide exam (http://www.mcasmentor.com/) uncovered several examples of ambiguous quantifier use, such as the following:

> Liu is separating the figures below according to their properties. There are at least three figures in each group. So far, he has made two different groups. List at least 3 figures that could go into each group. Explain what all the figures in each group have in common.

Among the ambiguities in the item is the instruction that [some] figures "could go into each group." Is that one group per figure, or a figure that could go into either group? Do all the figures have something in common regardless of group, or only by group?

Mathematics consultant Hill contributed this example from her teaching:

> Jalal has ten pockets and forty-four pennies. He wants to put his pennies in his pockets in such a way so each pocket contains a different number of pennies. Can he do it? Explain your answer.

The adult bilinguals in her elementary math class did not immediately grasp that "*each*" directed them to make 1 group per pocket. Nor did they understand without explicit explanation that they were to use all 44 pennies.

Consultant Grassetti is currently engaged locally in a study very related to the one we are proposing. The Amherst school district recently adopted *Investigations in Number, Space, and Time* (TERC, 1998)*,* which expressly integrates math concepts into a language-based curriculum. In a series of professional development days throughout the year, math teachers under Grassetti are evaluating the potential language difficulties, including ambiguities, for their mainstream and minority students.

Finally, pilot work with materials designed for this project has uncovered more and different non-target interpretations among the children than we had even anticipated. For example, an 11-year-old boy drew the sketch in Figure A to illustrate both "every boy is on a box" and "there is a box that every boy is on," whereas previous work (de Villiers & Roeper, 1993) led us to expect children older than 6 to distinguish the two.
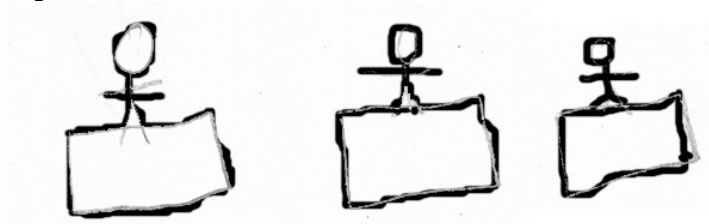
Figure A   Figure B

More unusual were the two drawings for the latter ("a box that every boy is on") from two children— one age 6;9 (Figure B), one age 14—with a box with a single boy on it. That interpretation appears to violate the presupposition that there is more than one (and probably more than two) exemplars required for *every*.

In a math probe which asked about the array in Figure C, "How many sets of 10 and how many ones are in the picture?"

Figure C.
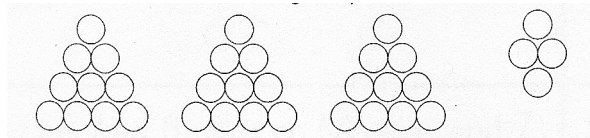
a girl aged 6;9 very carefully counted two sets of 10 to confirm that there were 10 in each of them, and then said "3 tens and 1 ones"—as if the question had an elided "sets": how many sets of ones were there.

Another 11-year-old gave a conceivable interpretation for "Do all the boxes have all circles?"—which he interpreted as "were they completely filled" with circles. The surprising part was that he later gave the same interpretation to "Does each box have each circle?" and "Does every box have every circle?", indicating that the different quantifiers did not have distinct semantic properties for him. Many of the children also unexpectedly required exhaustivity for the definite determiner in a "control" sentence: "Do boxes have circles inside?" answering "no" when two of three boxes had circles (and most adults respond "yes").

Such misinterpretations demonstrate the need for the current study to give a fuller picture of the behavior of *every* and related quantifiers in many different grammatical contexts within three major dialects.

D. Research Design and Methods:
i. Overview and design:
The proposed study has 3 major components: 1) a cross-sectional study of quantifier development in three dialects of English (MAE, AAE, and Spanish-influenced English); 2) statistical analyses (esp. correlation and regression) of the relation between the language probes and the math probes; and 3) a sketch of a causal relationship between elements of quantifier understanding and specific mathematics achievements.

There will be four phases of probes:

1) a dialect screener to confirm dialect group membership and to give a language score that does not involve quantifiers;
2) comprehension and production probes for quantifiers using abstract stimuli involving geometric shapes;
3) items with quantifiers in stories to explore how well children deploy their knowledge of quantifiers in real world contexts; and
4) mathematics probes, including word problems with and without quantifiers in them.

Following Orr (1987) and Mestre (1988), children will be prompted to tell why they have chosen the responses they do, and their explanations will be recorded, so the findings can be analyzed qualitatively as well as quantitatively (see "coding" and "analyses" below in Procedures).

Math probes will be the Beginning-of-the-year Math Assessments used by the Amherst Public Schools (based in part on the NSF- and other foundation-funded mathematics curricula, *Investigations* (TERC, 1998), for the primary grades, and *Math Connections*, for the older children (Herff Jones Education Division, 2005). These assessments give both a total score and a profile/ prescription based on which questions are answered correctly and which are missed.

ii. Participants:
The study will involve equal numbers of typically developing participants from the 3 dialect groups under examination, ranging from 4 to 14 (and adults). All participants will be matched for Educational Level or Parent Education Level as a proxy for socioeconomic status. There are well-known gender differences in math performance but we have no reason to believe that they will affect their quantifier use in probes of this type. Still, we will also attempt to include half male and half females, or at least balance the gender proportions across groups.

ii-a. The design is as follows;

Dialect by Age	MAE	AAE	ELL	Total
4	12	12	12	36
6	12	12	12	36
8	12	12	12	36

10	12	12	12	36
12	12	12	12	36
14	12	12	12	36
adult	12	12	12	36
Total	84	84	84	252

ii-b. Dialect status and general language score will be determined by the following procedures.

AAE and MAE speakers: The MAE- and AAE-speaking subjects will take the *DELV Screening Test (*Seymour, Roeper & de Villiers, 2003) to confirm the clinician's impressions of dialect and to insure that all subjects are typically developing. It will also yield a general language score that does not involve quantifiers. Part 1 assigns a category for Language Variation Status: "MAE," "some difference from MAE," or "strong difference from MAE." Part II gives a Diagnostic Risk score. The "strong difference" label is consistent with AAE, provided the child does not also show high risk for disorder. Part I is normed through age 12; part 2 through age 9. Therefore, for the older children and adults, we will have to supplement the screener with a language history questionnaire, and a clinician dialect rating scale, such as found in Wyatt (1995), and another general test, such as the *PPVT (*Dunn & Dunn, 1997). To get the densest examples of AAE dialects, the AAE subjects will be drawn from schools in urban working class neighborhoods (Washington & Craig, 1998). The other groups will be matched to the AA subjects.

Spanish-influenced English speakers: All the Hispanic children (or their parents) will complete a two-page language background questionnaire to establish where they fall on the bilingual spectrum (Genesee, Paradis, & Crago, 2004). Subjects need relatively good command of both English and Spanish. That is, one needs to look at relatively subtle semantic properties that would be overshadowed by much greater gaps in lexicon and grammar in the early stages of learning English, but on the other hand, the children also need to have enough competence in Spanish that it could be expected to have an influence on their English (Gathercole, 2002; Eilers, Pearson, & Cobo-Lewis, in press). In general, children will be selected who have had at least three years studying in the medium of English. Those who have not had 3 years of English-medium schooling, the 4- and 6-year-olds for the most part, can be given a part of the *Bilingual English Spanish Assessment*, or *BESA* (Pena, Iglesias, & Gutierrez-Clellen, forthcoming), a recently developed test normed on bilinguals which like the *DELV* can help characterize the subject's dialect and distinguish whether his or her language development is delayed relative to other bilinguals. The subjects too old for the *BESA* will take receptive vocabulary tests, the *PPVT-III* (Dunn & Dunn, 1997) and the *TVIP-H* (Dunn, Padilla, Lugo & Dunn, 1986) as an approximate measure of "language exposure." In our experience, the two scores are not completely comparable, but they are capable of making broad distinctions in performance (Pearson, Oller, Umbel, & Fernandez, 1996).

iii. The sites:

All subjects, including MAE speakers, will be drawn from schools in working class neighborhoods so that they can be compared to AAE speakers with the densest dialects possible. The research team enjoys a 15-year association with schools in stable working class neighborhoods in Hartford CT and Springfield MA where we can be confident of finding sufficient AAE-speaking and ELL subjects at all ages. Through the offices of the heads of Speech and Language Services in those public school districts we have been authorized to work in schools throughout the systems. (See letters attached).The adults will be drawn from undergraduate classes and members of minority-serving organizations on campus and neighboring communities.

iv. The materials/ stimuli:

Below are examples of pilot materials adapted from the methodology devised for the study of *some* and *all* reported in Roeper & Matthei (1975), updated in accordance with more current work. The materials test the semantic properties of exhaustivity, distributivity, scope/concord, and displacement (or movement) introduced above in the narrative. To the extent practical, all properties will be tested with all quantifiers: *every, each, most, some* and *all*. (Note that there is some overlap between the properties—for example, spreading is a kind of movement that affects how the quantifier will take scope. However, by collapsing them we would lose some distinctions that are conceptually or historically relevant, and we anticipate that at the exploratory stage the redundancy will be more helpful than harmful.) There are 3 types of quantifier

probes (Comprehension questions using abstract stimuli, Comprehension questions with a production (drawing) response, and Quantifier probes in text context) plus 1 set of mathematics probes.

a. Comprehension questions using abstract stimuli: The child is shown a set of figures and asked questions to probe their semantic interpretations. Their responses are followed up with probe questions: "why not?" or "show me." In pilot testing, we determined that children age 6 and up can do 50 such items in approximately 15 minutes. The 4-year-old version will be somewhat abridged. Examples of items for each concept are below. (Note that the sequences of related questions are not given in a batch.

Items to test EXHAUSTIVITY: Some of the terms require that we exhaustively examine a set, i.e. name all its members. They stand in contrast to a plural (in b) which does not have this requirement for adults:

Figure 1.
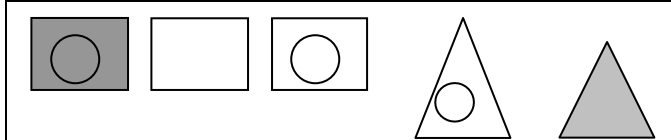
Example questions for Figure 1:
a. Do all the boxes have a circle inside => exhaustive (no)
b. Do boxes have a circle inside => non-exhaustive (yes)
c. What has a circle=> exhaustive (point to 2 boxes and 1 triangle)

Items to test **DISTRIBUTIVITY** / **COLLECTIVITY:** Whether the items are considered singly (distributed) or as a group (collectively). Which words allow both readings? Which require distribution?
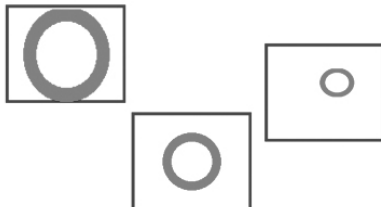
Figure 2.

Example questions for Figure 2:
f. Do the boxes have three triangle tops=> collective (yes)
g. Do all the boxes have three triangle tops => collective (yes)/ distrib (no)
h. Do most of the boxes have three triangle tops => no
i. Does every box have three triangle tops => no

To test CONCORD/ SCOPE: Are 2 instances of a quantifier interpreted as 1 quantification (as in multiple negation) or as 2 (as in MAE negation)? Is one quantifier interpreted in terms of another? Is the second quantifier distributed with respect to the items separately or as a group?

Figure 3.:

Example questions for Figure 3:
n1. Do all the boxes have a circle => yes
n2. Do all the boxes have all circles => if yes = concord/ if no, non-concord.

To test SCOPE: Example questions:

Show a picture of 4 boys and 3 balloons, and ask:
> Is every boy holding a balloon?   No (not a problem after age 5)

Show 3 boys and 4 balloons (classic spreading example)
> Is every boy holding a balloon? No, not that balloon.

Show 3 boys and 3 balloons (1 each)
> Is every boy holding every balloon?
> > Yes, one quantification
> > No, one in scope of the other
> > (would need strings for each boy to hold all the balloons)

To test MOVEMENT/ DEPLACEMENT: Some quantifiers apply to nouns they are not next to. The quantifier may move away from its noun; eg. may be copied (as with spreading) in another part of the sentence.
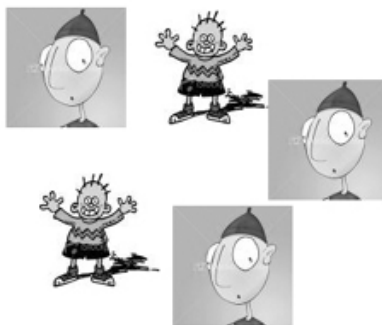Set-up 1 as in Figure 3.

Example questions:
> r1.   Does every box have a circle => yes
> r2   Does a box have every circle => no

Figure 4.

Example questions for Figure 4: "Every boy does not have a hat? Is that right?   Show me."
> Answers Yes If: = (not every boy) has a hat --will point to the boys with no hats.
> Answers No: If: =   every boy (does not have a hat) – will point to the boys with hats.

b. Production Questions
Children will also be asked to show their understanding by drawing with crayons what the sentence says. In piloting, children age 8 and up could do 12 of these in less than 10 minutes; 6-year-olds took longer (with laborious drawings); and 4-year-olds often discontinued, (e.g. drew a boy "who wanted to be on the grass" instead of a box) so their production data will have to be treated separately.

Ex. Draw a picture with lots of circles
- Where all the circles are black.
- Where all the circles are all black.
- Where the circles are some black.
- Where every circle has no black.

Draw a picture with 3 boys:
- Where every boy is on a box.
- Where there is a box that every boy is on.

Draw me some flowers and vases, like this:
- The flowers are all in a vase.
- All of the flowers are in vases.
- Each flower is in a vase.
- Each flower is in vases.

- The flowers are each in a vase.
- Every flower is in every vase.

c. Quantifier probes in text context (with real world objects)
In piloting, these items were added in with the math probes to insure that all diagnostic tests had a minumum of 10 word problems, with some devoted to each semantic property. The "enriched" math probes took approximately 20 minutes.

EXHAUSTIVITY
- Show picture of 8 children, 6 of them eating a cookie. Show me: Who is eating a cookie? (Will the child point to one (non-exhaustive)? or 6 (exhaustive)?
- John picks up his hat and coat. Did John pick up the coat? (yes) Was it the coat John picked up? (no, here the cleft forces an exhaustive answer: he picked up the <u>hat and coat</u>.)

DISTRIBUTIVE/ COLLECTIVE
- Four children gave one flower to three teachers and three aides. How many flowers were given?
  - If "three teachers" equals "each of three teachers" and if "four children" = "each of 4 children" = 24.
  - If "three teachers" is one group of three teachers and "three aides" one group of 3 aides = 8.
  - If group of four children gave to teachers and aides, one each = 6.
- Who got the largest piece of chocolate cake? Compare:
  - John has 1/10 and Bill has 1/4 and Fred has 1/3. (This seems to imply a single cake at the outset.)
  - John has 1/10 of a cake, Bill has 1/4 of a cake, and Fred has 1/3 of a cake (This is more likely to be separate cakes and need not sum to less than one.)

v. Procedure: Each child will be tested individually in at least two sessions. During the first session the child will take the screening tests to determine dialect and get a general language score. They will do the first set of abstract stimulus probes. During the second session (and possibly third for the younger children) they will do the math diagnostic test and the production probes and word problems. The experimental items will be in quasi-random order (with two versions, so that order effects can be evaluated). The comprehension items will be administered with a tablet PC with a touch screen and built in recorder and camera; the production items will be done in a small booklet (which will also include the student assent form).

The think-alouds, either while the child is drawing or in response to prompting from the examiner, will be audio recorded for later transcription, with a video record to aid transcription.

vi. Coding and analysis:
Coding of Quantitative analyses:
> The answer to each item will be coded as to whether it agrees with adult-MAE or not, and further whether the subject's answer is distributive, collective, exhaustive, displaced, scopally different, off numerically, or just plain "off-base." (Note: we will use "adult-MAE" as the reference point, because distance from it is hypothesized to create non-mathematical difficulties with math problems stated in MAE. Its use as a reference point does not imply that it is a "standard" for the other dialects. In fact, each dialect group will have adult controls to establish standard targets within the dialect.) For the math scores, there will be a total score of correct answers and a profile of concepts taken from the diagnostic: e.g. knows numbers to 10, knows numbers to 100, one-to-one counting to 5; to 10; to 100; ordinality, sequencing, age-appropriate pattern recognition, basic adding; basic subtracting, comparison of fractions, addition of unlike fractions, etc.

Coding of Qualitative analyses:
> Student "think-aloud" responses will be evaluated for 1) typicality, 2) agreement with the coding for quantitative analyses, and 3) explanatory value. Responses with high explanatory value will be transcribed and used for discussion as well as confirmation of the quantitative findings.

Analyses

1. descriptive statistics: For each of the groups in the research design, basic descriptive statistics (mean, median, mode, standard deviation, kurtosis) will be prepared and presented in tabular and graphical form for all variables (e.g. total agreement to adult target within dialect, agreement to MAE-target; target answers by lexical item; percent of distributive answers, of exhaustive answers in obligatory contexts, of collective answers, displaced etc., overall and by lexical item). Each variable will be examined for outliers and distributional characteristics.

2. Specific Analyses for the Research Questions (RQ):

RQ1-a. What are the steps and approximate timetable for the acquisition of the semantic properties of the different quantifiers for native-MAE speaking children?

> Analyses of Variance and chi-square analyses will focus on the 84 native MAE subjects and will display tables and graphs by age of the following variables: for the quantifiers, percent of "agreement to adult" for each semantic property and each quantifier word both within semantic property and in general. The "agreement" variable coded for the think-aloud answers (agreement with the short answer given) will be compared to the agreement-to-adult scores. For example, if the four-year-olds are shown not to spread, the think-aloud data can be examined to confirm which elements in the pictures or stories the children were attending to and to better determine whether their answers are on target or not. The two sets of data should provide the basis for describing the broad outline of development of the interpretation of the semantic properties in general and in conjunction with specific lexical items for native-MAE speaking children.

RQ2. Are the developmental stages in understanding quantifiers (in RQ1) reflected in children's math skill? Is understanding of the semantic properties of quantifiers correlated with success in math generally, and with specific mathematics concepts?

> Quantitative analyses: 1) Math diagnostic scores will be correlated (with Age partialled out) to an "MAE-target" score for the quantifiers; math scores will also be correlated to a general language score derived from the *DELV Screener* Part 2, or *PPVT/ TVIP*. 2) Regression analyses will be done with quantifier scores and general language scores as independent variables (along with gender, age, parent education level etc.). Having two language scores will permit testing whether the quantifier scores make a specific contribution to math performance beyond what is explained by variations in language skill more generally.
>
> Qualitative analyses: Math diagnostic profiles will be aligned with quantifier profiles in contingency tables. For example, how many children who do not succeed on statements of inequalities in math fail on the "most" quantifier items, compared to children who do succeed on statements of inequalities? What percentage of distributive answers do children who have not mastered one-to-one counting give compared to children who have mastered one-to-one counting?

RQ3: Are the learning patterns (the timetables in RQ1a and b) different for children from different dialect communities?

> This question will be addressed by the Analysis of Variance (Type III general linear model) similar to that in RQ1. The first consideration will be to compare the adult groups to test the status of any differences observed by dialect group. The full MANOVA for different dependent variables will show where group differences and important interactions are found, using the adult MAE group as the "standard" (as in Coding above). A descriptive analysis of the path of acquisition, as described under RQ1 for MAE, can also be done using the adult groups for the other two dialects as the standard for each group (distributively!)

RQ4: Are the associations between the stages of quantifier learning and mathematics achievement (RQ2) different for children speaking native-MAE versus AAE versus Spanish-influenced English?

> The correlation, regression, and contingency analyses described under RQ2 for the native MAE-speakers will be performed by dialect group and for the whole group combined (with dialect as an additional factor for the regression).

Database:
Verbatim responses and codes will be stored in a relational database, with summary files organized by child, by lexical item, by semantic property and by math level. (This will be similar to the ones for the *DELV* and the BSG projects, both involving multiple measures from over 1000 subjects.) The extensive data generated by the study will be appropriate for large numbers of exploratory studies. The combination of qualitative and quantitative data will provide greater confidence in the findings than could be provided by either type of data alone. The analyses for the four research questions will identify vulnerabilities for each dialect group which can then be tested for their relevance to math performance (in a further study).

vii. Testing schedule and staff allocation
We anticipate that older children will be able to complete all 4 probes in two sessions, but that younger children may take a third session for them. We also anticipate that we will not be able to use all the subjects who take the screening tests, as some will not have a clear dialect status or some ELLs will not show adequate Spanish (or possibly English) ability to be included. Therefore we will recruit at least 20% more potential subjects than are included in the design. Total sessions will equal 120% of 252=302 first sessions, 100% of 252 second sessions, and 60% of 252 third sessions=151 third sessions, or a total of about 705. Initial testing will target adults, then children ages 6, 10 and 14 to insure data across the years of school instruction; early and intermediate ages (4, 8, and 12) will be filled in during year 2.

On the basis of previous experience with testing students in the local public schools with staff from the University of Massachusetts, it is estimated that the average tester can do 5 sessions in one day during approximately 100 school days available per year (taking out the first month of school, holidays, testing days, etc.) or 500 sessions per year and 750 during the first year and a half of the project. The work can be done with one 20-hour tester (4 five-hour days/wk) and additional graduate student and PI testers if needed.

Timeline (24 months):
Yr 1: Month1-convene advisory board; assemble equipment and materials; re-initiate contacts with schools and testing sites get parent consents (IRB approval is in process now, before the grant period); Month2-begin testing counting 5 spring semester months with at least 60 sessions/month (ss/mo), 3 summer months at 25 ss/mo, 3 fall semester months at 60 ss/mo totaling 550 sessions in the first year (months 2 – 12). Summer months will also be used to consolidate databases, code data thus far, and do pilot analyses.
Yr. 2: Month 13 re-convene advisory board, and within Months 13 to 18, there will be another 5 spring semester months with 60 ss/mo or 300 more sessions. Month 16: attend (and hopefully present at) NCTM meeting in Salt Lake City; Month 17: attend/ present at Int'l Child Language meeting (Edinburgh); Months 19-21: Final data preparation and analysis. Goals for months 22-24: Month 23 Boston Language Development Conference, Month 24.2 MAA meeting in DC (Jan 3, 2009); workshops with local and regional mathematics faculty and teacher training programs; at least one article submitted to linguistics journal, one to mathematics education publication.

viii. Data-sharing. Findings will be disseminated through the usual channels of technical reports, conference presentations, and journal articles as indicated in the Timeline. The PIs and their linguistics advisors are regular participants in child language and language disorder conferences. Several of the mathematics advisors are regular participants (or for M. Pearson, organizers of) the math education conferences and workshops so results can be shared in their conferences and journals as well.

In addition, steps will be taken to make the raw data (without personal indentification information) available to the research community. The child language research community is fortunate to have several excellent vehicles for data sharing. The co-PI has one corpus on CHILDES, the Child Language Data Exchange System (MacWhinney, 2000) and is currently converting the archived materials from the Working Groups on AAE into a flexible text format that can be input into any one of a number of computerized analysis systems such as SALT and CP. Although transcripts from this project will not be typical language samples), we expect that they will be of great interest to the math education community.

LITERATURE CITED:

Bach, E., Jelinek, E., Kratzer, A., & Partee, B. H. (Eds.) (1995). *Quantification in Natural Languages*, Dordrecht: Kluwer.

Brookes, P. & Braine, M. (1996). What do children know about the universal quantifiers "all" and "each"? *Cognition, 60*, 235-268.

Craig, H. K. (1996). The challenges of conducting language research with African American children. In A. G. Kamhi, K. E. Pollock, and J. L. Harris (Eds.), *Communication Development and Disorders in African American Children: Research, Assessment, and Intervention* (pp. 1-18). Baltimore: Brookes.

Craig, H. K. & Washington, J. A. (2000). An assessment battery for identifying language impairments in African American children. *Journal of Speech, Language, and Hearing Research, 43*, 366-379.

Crain, S. (2000). Sense and sense ability in child language. In S. C. Howell, S. A. Fish & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University conference on language development* (pp. 22–44). Somerville, MA: Cascadilla Press.

Crain, S., Thornton, R., Bosta, C., Lillo-Martin, D., & Woodams, E. (1996). Quantification without qualification. *Language Acquisition 5*(2), 83-153.

DellaCarpini, M. (2003). "Developmental stages in the semantic acquisition of quantification by adult L2 learners of English: A pilot study". In J.M. Liceras, H. Zobl and H. Goodlulck (eds.), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002): L2 Links*. Somerville, MA: Cascadilla Press.

Denton, K. & West, J. (2002). Children's reading and mathematics achievement in kindergarten and first grade. *Education Statistics Quarterly, 4 (1),* 19-26.

de Villiers, J. & Roeper, T. (1993). The emergence of bound variable structures. In E. Reuland & W. Abraham (Eds.), *Knowledge and Language: From Orwell's Problem to Plato's Problem.* Dordrecht: Kluwer.

Drozd, K. F. (2001). Children's weak interpretation of universally quantified sentences. In Bowerman, M. and S. Levinson (Eds.) *Language Acquisition and Conceptual Development*, (pp. 340-376). Cambridge: Cambridge University Press.

Dunn, L.M. & Dunn, L. (1997). *Peabody Picture Vocabulary Scale – Third Edition* (PPVT-III). Circle Pines, Minnesota: American Guidance.

Dunn, L. M., Lugo, D., Padilla, E. R. & Dunn, L. M. (1986). *Test de Vocabulario en Imagenes (TVIP).* Circle Pines, Minnesota: American Guidance.

Eilers, R. E., Pearson, B. Z. & Cobo-Lewis, A. B. (in press). The social circumstances of childhood bilingualism: The Miami experience. In E. Hoff & P. McCardle (Eds.), *Childhood bilingualism.*Clevedon UK: Multilingual Matters.

Gathercole, V. C. (1997). The linguistic mass/count distinction as an indicator of referent categorization in monolingual and bilingual children. *Child Development, 68* (5), 8332-842.

Gathercole, V. C. (2002). Command of the mass/count distinction in monolingual and bilingual children: An English morphosyntactic distinction. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy development in bilingual children* (pp. 175-206). Clevedon, UK: Multilingual Matters*.*

Gelman, R. & Gallistiel, C.R. (2004, Oct. 15) Language and the origin of numerical concepts. *Science*, 306, 441-443.

Genesee, F., Paradis, J. & Crago, M. (2004). *Dual-language development and disorders.* Baltimore: Paul Brookes.

Guerts, B. (2002). *Quantifying kids.* Ms. (downloadable at www.phil.kun.nl/tfl/bart).

Herff Jones Education Division, (2005). *Math connections.* Armonk, NY: About Time Press.

Hirigoyen, H. (1997). Dialectal variation in the language of mathematics: A source for multicultural experience. In J. Trentacosta & M. Kenney (Eds.), *Multicultural and gender equity in the mathematics classroom: The gift of diversity* (pp. 164-168). Reston, VA: National Council of Teachers of Mathematics (pub.)

Inhelder, B. and J. Piaget (1964). *The Early Growth of Logic in the Child.* London: Routledge, Kegan and Paul.

Khisty, L. L. (1995). Making inequality: Issues of language and meaning in math teaching with Hispanic students. In Walter G. Secada, E. Fennema, & L. B. Adijian (Eds.), *New agendas for equity in mathematics education* (pp. 279-97). NY: Cambridge University Press.

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk." *Developmental Psychology, 42 (1),* 59-69.

Kramer, I. (2000). *Interpreting indefinites: an experimental study of children's language comprehension.* Doctoral dissertation, MPI Series in Psycholinguistics.

Labov, W. (1972). *Language in the inner city.* Philadelphia PA: University of Pennsylvania Press.

Lakoff, G. & Nunez, R. E. (2000). *Where mathematics comes from.* New York NY: Basic Books.

Lidz, J. & Musolino, J. (2002). Children's command of quantification. *Cognition, 84*, 113-154.

MacWhinney, B. 2000. *The CHILDES project: Tools for analyzing talk. Third Edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Mestre, J. 1988. The role of language comprehension in math and problem solving. In R. Cocking & J. Mestre (Eds.), *Linguistic and cultural influences on learning math* (pp. 201-220). Hillsdale, NJ: Erlbaum.

Mestre, J., Gerace, W. & Lochhead, J. (1982). The interdependence of language and translational math skills among bilingual Hispanic engineering students. *Journal of Research in Science Teaching, 19*, 339-410.

Mestre, J., Thibodeau, P. & Gerace, W. (1988). Comprehending premises: Effects of negations and training among Anglos and Hispanics. *NABE Journal, 12 (3),* 243-279.

Murray, M. (2004). *Teaching mathematics vocabulary in context.* Portsmouth NH: Heinemann.

Nishigauchi, T. (1999). Some preliminary thoughts on the acquisition of the syntax and semantics of *wh-*constructions. In *Theoretical and Applied Linguistics at Kobe Shoin 2.* Kobe Shoin Institute for Linguistic Sciences.

Oller, D. K. & Eilers, R. E. (Eds.) (2002a). *Language and literacy development in bilingual children.* Clevedon, UK: Multilingual Matters.

Oller, D. K. & Eilers, R. E. (Eds.) (2002b). An integrated approach to evaluating effects of bilingualism in Miami school children: The study design. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy development in bilingual children* (pp.22-42). Clevedon, UK: Multilingual Matters.

Orr. E. W. (1987). *Twice as less: Black English and the performance of Black students in mathematics and science.* NY, NY: Norton.

Papafragou, A. & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics-pragmatics interface. *Cognition, 84,* 253-282.

Pearson, B. Z. (1993). Predictive validity of the Scholastic Aptitude Test for Hispanic bilingual students. *Hispanic Journal of the Behavioral Sciences, 15,* 342-356.

Pearson, B. Z. (2002). Narrative competence in bilingual school children in Miami. In D.K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 135-174). Clevedon, UK: Multilingual Matters.

Pearson, B. Z., Fernandez, S., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning, 43,* 93-120.

Pearson, B. Z., Oller, D. K., Umbel, V. M., & Fernandez, M. C. (October, 1996). *The relation of lexical knowledge to measures of literacy and narrative discourse in monolingual and bilingual children.* Paper presented at the Second Language Research Forum, Tucson, AZ.

Pena, E., Iglesias, A., & Guitierrez-Clellen, V. (in press). *Bilingual English/Spanish Assessment, BESA.* San Antonio, TX: Harcourt Assessments.

Penner, Z. (2006). *Korrelationen zwischen Mathetest und Sprachtest.* (ms.)

Pérez-Leroux, A. T. & Roeper, T.(1999). Scope and the structure of bare nominals: evidence from child language. *Linguistics 37,* 927-960..

Philip, W. (1995). *Event quantification in the acquisition of universal quantification.* Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Philip, W. (1996). Symmetrical interpretation and scope ambiguity in child Dutch. In C. Koster and F. Wijnen (Eds.), *Proceedings of GALA 1995.* Gröningen, the Netherlands.

Philip, W. 2004. "Two theories of exhaustive pairing." Ms., Utrecht Institute of Linguistics, OTS. Utrecht University, The Netherlands.

Riley, M. S. & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction, 5* (1), 49-101.

Roeper, T. (in press). *The prism of language*. MIT Press.

Roeper, T. & Matthei, E. (1975). On the acquisition of *some* and *all*. Presented at the Sixth Child Language Research Forum, Stanford University, April 1974. Appeared in *Papers and reports on child language development* (1975), Stanford University, 63-74.

Roeper, T., Pearson, B. Z., Penner, Z. & Schultz, P. (July 2002). *The emergence of wh-variables: Cross-linguistic explorations.* Paper presented at the IX International Congress for the Study of Child Language (IASCL) and the Symposium on Research in Child Language Disorders (SRCLD), University of Wisconsin, Madison, WI.

Roeper, T., Schulz, P., Pearson, B. Z. & Reckling, I. (2005). From singleton to exhaustive: The acquisition of *wh-*. In *Proceedings of SULA* (Semantics of Understudied Languages), Amherst MA.

Roeper, T., Strauss, U. & Pearson, B. Z. (2006). The acquisition path of quantifiers: Two kinds of spreading. In T. Vanjivik (Guest Editor), University of Massachusetts Occasional Papers. Special Issue on *Language Acquisition.*

Seymour, H. N. & Pearson, B. Z. (Eds.) (February 2004). Evaluating language variation: Distinguishing dialect and development from disorder. Special issue of *Seminars in Speech and Language*, 25 (1)..

Seymour, H. N., Roeper, T. & de Villiers, J. G. (2003). *Diagnostic Evaluation of Language Variation (DELV) Screening Test.* San Antonio, TX: The Psychological Corporation.

Seymour, H.N., Roeper, T. & de Villiers, J.G. (2005). *Diagnostic Evaluation of Language Variation (DELV) Norm-Referenced.* San Antonio, TX: Harcourt Assessments.

Seymour, H. N., Roeper, T. de Villiers, J.G., de Villiers, P.A. & Pearson, B.Z. (April, 2002). *Developmental Milestones in African American English Child Speech: Report to NIH.* Ms., Department of Communication Disorders, University of Massachusetts, Amherst MA.

Seymour, H. N. & Seymour, C. M. (1977). A therapeutic model for communicative disorders among Black English speaking children. *Journal of Speech and Hearing Disorders, 46*, 274-280.

Spanos, G., Rhodes, N. C., Dale, T. C. & Crandall, J. (1988). Linguistic features of mathematical problem solving. In R. Cocking & J. Mestre (Eds.) *Linguistic and cultural influences on learning mathematics.* (pp. 221-240). Hillsdale, NJ: Lawrence Erlbaum.

Strauss, U., Roeper, T. Pearson, B. de Villiers, J. & Seymour, H. N. (2003). *Acquisition of exhaustivity in wh-questions.* Paper presented at the UMass Workshop on Acquisition of Quantification, May 20, 2003.

Takahashi, M. (1991). Children's interpretation of sentences containing *every*. In T.L. Maxfield and B. Plunkett (Eds.), *Papers in the acquisition of wh*. Amherst, MA: GLSA Publications.

TERC, Author. (1998). *Investigations in Number, Data, and Space.* White Plains, NY: Dale Seymour Publication, White Plains, NY

Tunstall, S. (1998). *The interpretation of quantifiers: Semantics and processing*. Doctoral dissertation, University of Massachusetts.

Varley, R., Klessinger, N., Romanowski, C., & Siegal, M. (2005). Agrammatic but numerate. *Proceedings of the National Academy of Sciences*, 102, 3519-3524.

Washington, J. A. & Craig, H. K. (1998). Socioeconomic status and gender influences on children's dialectal variations. *Journal of Speech, Language, and Hearing Research*, *41*, 618-626.

Wolfram, W. & Fasold, R. W. (1974). *The study of social dialects in American English.* Englewood Cliffs NJ: Prentice-Hall.

Wyatt, T. (1995) Wyatt, T. (1995). Language development in African American English child speech. *Linguistics and Education*, 7, 7-22.