

UNBIASED ASSESSMENT OF FIRST LANGUAGE ACQUISITION IN ENGLISH: DISTINGUISHING DEVELOPMENT AND DIALECT FROM DISORDER

Peter A. de Villiers and Jill G. de Villiers, Smith College

Thomas Roeper, Harry N. Seymour, and Barbara Zurer Pearson, University of
Massachusetts, Amherst

1. INTRODUCTION

In the United States, as in all countries of the world in which English is widely spoken, there exist different dialects or variants of the language. These dialects are often defined by regional or cultural groups and may vary from each other in one or more of several aspects of the language – phonology, morphology, syntax, lexical semantics, or pragmatics (ASHA 2003, Wolfram 1991). Thus, we can distinguish between Cajun English, spoken in much of Louisiana, Appalachian English, spoken in the states along the Appalachian mountains, especially Kentucky and West Virginia, or the broader Southern American English spoken across the southeastern states. All of these dialects reflect coherent rule-governed varieties of the English language.

Many, though not all, African Americans speak a distinctive variety of English called "African American English" (AAE) (Labov 1972, Mufwene 1998, Rickford 1999). AAE is far less geographically defined than the other dialects of American English mentioned above, although there is some regional variation. Rather, AAE emerged from the social separation of racial segregation (and before that of slavery) as the form of English spoken by a culturally defined group of people. Children or adults of other races who have strong cultural identification or primary social interaction with African Americans may also speak AAE. Thus, AAE may be defined by a set of characteristic linguistic features that distinguish it from mainstream forms of American English (Green 2002).

For example, aspect has priority over tense in AAE. Thus, there are characteristic markers of aspect as in the sentences *He **be** workin' all the time* (indicating habitual action) and *She **bin***

done ate that chicken (indicating completion of action). On the other hand, simple past tense and third person present tense inflections are frequently omitted, as in *Yesterday he cook a turkey* and *She like to sit in that chair*. Similarly, the copula forms of *BE* are only obligatory in AAE in sentence contexts in which they carry meaning. So the present tense copula is usually omitted in utterances such as *He tall*, but it is required in the dialect and is always provided by adult speakers in past tense contexts such as *Last night she was real angry*. Like other dialects of English such as Belfast English (Northern Ireland) and Southern American English, AAE uses negative concord, as in *He don't have no overcoat* and *I ain't gonna eat no broccoli*. Several distinctive phonological processes include the deletion of some final consonants and the reduction of consonant clusters, but only in final position in a word. So an AAE speaker will say *Las' year I had to take a tes'*, but *Yesterday I stood in line for an hour*. In final position [f] is substituted for voiceless *th* [θ] (*baf* for *bath*, but not *fin* for *thin*). Similarly, [v] is substituted for voiced *th* [ð] (*breave* for *breath*), but again, only in the final position. In initial position, [d] will often be produced in place of [ð] (*den* for *then*)¹.

Over the past twenty years, the American Speech-Language-Hearing Association (ASHA 1983) has acknowledged that there are problems in the accurate language assessment of children who come from different cultural backgrounds and speak dialects of English differing from Mainstream American English (MAE)². The existing standardised tests developed in the USA to establish whether a child has a language disorder or delay take MAE as the norm. Furthermore, the most widely used tests choose those aspects of MAE most easily measured and understood by the speech-language pathologist. Thus, the focus of major parts of these tests is on assessing acquired vocabulary (especially labels for pictured objects that the child must provide or comprehend) and grammatical morphemes in obligatory linguistic contexts (such as plural *-s*, past tense *-ed*, possessive *-s*, third person present tense *-s*, and copula *BE*). Additional reasons for concentrating on grammatical morphemes in assessment include empirical research and theoretical arguments that morphosyntax and other functional elements of English may be particularly vulnerable in language disorders (Leonard 1998; Rice and Wexler 1996, 2001).

However, acquired vocabulary, especially names for objects, is particularly subject to cultural as well as linguistic variation across dialects of English, dependent as it is on exposure to

words for objects in particular domains of experience (e.g., names for animals, foodstuffs, tools, etc.). Even worse, focus on the morphosyntax of MAE happens to test those areas of the grammar of English in which there is considerable dialectal variation, and in which AAE, in particular, differs from MAE. Children with language impairment frequently omit tense markers and other grammatical morphemes for an extended period of time, but these morphemes are also omitted in AAE as part of the normal rule system of the dialect. Thus, if the procedures and norms of these MAE language tests are used with African American children, it becomes difficult to tell whether a child scores low on the test because he is a normally developing AAE dialect speaker, or because he has a language disorder or delay (Wyatt 2002). In the light of this problem, most of the available tests include a warning about use with dialect speakers and even identify items that should be credited or omitted for such speakers. However, this raises difficulties in the use of the norms for those parts of the tests. Not surprising then, African American children are over-represented in the caseloads of speech-language pathologists and especially in special education classes, where eligibility is often determined on the basis of vocabulary and other language assessments. It seems likely to us that many of these children are misdiagnosed as language-impaired because of the mismatch between their dialect and the language assessment instruments employed. On the other hand, it is possible that other African American children who do have genuine language impairments are not receiving language intervention services because the deviations in their language from the mainstream patterns is simply "assumed" to be dialectal³.

It is in this context that we have been collaborating since 1998 on a research contract from the US National Institute on Deafness and Communicative Disorders (NIDCD)⁴ to develop and standardise a dialect-neutral language assessment instrument for African American children aged 4 through 9 years. The collaborating group of researchers and scholars includes linguists, developmental psychologists, researchers on communication disorders, and speech-language pathologists⁵. The project thus reflects the combined expertise of linguists studying the distinctive characteristics of AAE, developmental psycholinguists with extensive experience studying normal acquisition in several domains of language, and clinicians skilled in assessment of and intervention with language-impaired children. Another important feature of the project was that the research contract included an agreement between the principal investigators, the federal funding agency, and a major publishing company (The

Psychological Corporation) for the publishers to carry out the field testing and standardisation of the assessment instrument as soon as the researchers developed the materials. This considerably reduced the time from conception to publication of the tests.

The overall goals of the project were to develop a comprehensive language assessment of syntax, semantics, pragmatics, and phonology for children between 4 years and 10 years old. The test should be able to determine whether the language variation observed in the children was due to normal development with age, dialect differences, or language impairment. Most importantly, the diagnostic parts of the test should not be biased against dialect speakers, especially children acquiring AAE. This was a rather tall order, and it was with some trepidation that we set out on the project. Over the course of the project, specific strategies emerged to fulfill these goals. First, the testing procedures distinguished between **contrastive** items that systematically vary between AAE and MAE and **non-contrastive** items in which both AAE and MAE require the same response from the child (Seymour and Seymour 1977). Contrastive items include various aspects of the morphosyntax and phonology of English. These items are only used for identifying deviation of the child's language from MAE towards AAE dialect, not for diagnosing possible language impairment. Non-contrastive items include morphosyntactic features obligatory in both MAE and AAE (e.g., past tense copula *was/were*) as well as a variety of other fundamental aspects of syntax that are common across different dialects of English. In this respect, the goal was to assess syntactic knowledge at a deeper level of sentence structure than the surface morphosyntax of the language. We will illustrate one of these areas, complex wh-question comprehension, later in this paper. Similarly, the semantic assessment avoided focusing on acquired vocabulary and attempted to assess both the child's lexical organisation (relationships between words in their vocabulary) and the process of vocabulary learning. This, too, we will illustrate with an example later in the paper. Finally, the pragmatics assessment focused on functional elements of language and communication that are essential for children's success in early schooling, including speech act understanding, question asking, and narrative skills. We believe that the strategies adopted in our project as well as many of the actual assessment procedures created for the test are particularly relevant for the development of accurate and effective assessment instruments in any situation in which children acquire a variety of dialects of English that vary from what

has been held to be the "standard" or received form. This is clearly the case in the South African setting (Mesthrie 2002).

In the remainder of the paper, we will describe the characteristics of the assessment instrument that emerged from our research project and highlight testing procedures from three particular subtests, one each from syntax, pragmatics, and lexical semantics, that illustrate the principles and strategies adopted to create a dialect-neutral assessment.

2. THE DIAGNOSTIC EVALUATION OF LANGUAGE VARIATION

The Diagnostic Evaluation of Language Variation (DELV) has two versions, namely a screener (DELV-SC) that takes 15 to 20 minutes to administer, and a separate, more comprehensive test that takes 45 minutes to administer and is used for the primary diagnosis of language problems in children aged 4 through 9 years. The longer test is already available in a criterion-referenced form (DELV-CR) and is currently being norm-referenced on both a substantial sample of African American children, and a sample of children reflecting the fuller US population.

2.1 The DELV Screener

The DELV-SC contains a variety of morphosyntactic and phonological items on which AAE-speaking children produce systematically different responses from speakers of MAE. These contrastive items are used only to identify the children as speakers of a dialect that differs from MAE, not for diagnostic purposes. The morphosyntax dialect **identifier** items include the production of third person present tense *-s*, the production of *don't* versus *doesn't* and *have* versus *has*, and present tense *BE* copula forms. Examples of distinctive AAE phonology tested on the DELV-SC are the substitution of *baf* for *bath* and *gif* for *gift*. Standardisation of the DELV-SC on a large sample of over 1000 4- through 9-year-olds demonstrated that children growing up in AAE-speaking African American communities and characterised by clinicians as AAE dialect speakers were reliably distinguished from MAE-speaking children by their scores on these items.

The DELV-SC also includes sets of **diagnostic** items designed to tell the clinician whether more comprehensive testing is necessary because the child is at risk for language delay or impairment. The diagnostic items are non-contrastive between AAE and MAE dialects and are based on three influential theories about the nature of specific language impairment (SLI) in children. One of these suggests that problems with morphosyntax are fundamental in SLI (Leonard 1998; Rice and Wexler 1996). The diagnostic part of the DELV-SC therefore includes production of past tense *was/were* copula forms (obligatory in both AAE and MAE) and possessive marking on pronouns (also obligatory in both AAE and MAE). A second theory of the nature of SLI highlights the children's difficulties in processing incoming speech, as tested by non-word repetition tasks (Campbell, Dollaghan, Needleman, and Janosky 1997; Dollaghan and Campbell 1998). Thus, the DELV-SC contains six non-word repetition items that vary in number of syllables and only use phonological elements that are common to both AAE and MAE. Finally, other researchers have focused on the difficulties that some language-impaired children have with syntactic forms that involve movement rules (Van der Lely 2003), and the DELV-SC tests children's comprehension of simple and complex *wh*-question forms that depend on *wh*-movement. Standardisation of the DELV-SC revealed strong developmental growth on these diagnostic items between the ages of 4 and 9 years. Children at these ages who were *a priori* classified by experienced speech-language pathologists as language-impaired made significantly more errors on the items than typically developing children of the same age. However, there were no differences in performance on the diagnostic items between AAE- and MAE-speaking children at any of the ages.

The DELV-SC therefore provides the clinician with an assessment of the child's **language variation status**, indicating whether the child speaks MAE, "some variation from MAE", or "strong variation from MAE". It also provides a dialect-neutral assessment (at least for AAE, MAE, and similar dialects) of the child's **diagnostic risk status**. This is categorised as: lowest risk for language disorder (no follow-up testing needed); low to medium risk (follow-up at the clinician's discretion); medium to high risk (comprehensive follow-up testing recommended); or highest risk (comprehensive follow-up testing recommended).

2.2 The DELV Comprehensive Test

The comprehensive diagnostic assessment instrument has 11 subtest components organised into four general language domains, namely syntax, pragmatics, semantics, and phonology. In each of these subtests, only non-contrastive language features are assessed. That is, the DELV-CR focuses on aspects of English for which the different dialect speakers produce the same forms for the same functions. The subtest components are briefly described in Table 1.

Table 1. Subtest Components of the DELV-CR

Subtest	Aspect of Language Assessed
Wh-question Comprehension	Knowledge of syntactic variables, movement rules, syntactic barriers to movement, embedded clauses
Wh-question Production	Asking the right question to discover missing information
Passive Sentence Comprehension	Knowledge of movement rules and implied agents
Communicative Role-Taking	Understanding the point of view of speakers and the speech acts they are producing
Short Narrative Production	Reference specification, expressions of temporal relationships between events, understanding the mental states of the protagonists (theory of mind)
Article Production <i>a</i> vs <i>the</i>	Understanding of varieties of meaning expressed by articles in English
Verb Contrast Production	Organisation of the child's verb lexicon into contrasting words and levels of meaning
Preposition Contrast Production	Organisation of the child's prepositions into contrasting terms
Quantifier Comprehension	Understanding the meaning and scope of the quantifier <i>every</i>
Syntactic Bootstrapping of Verbs	Fast mapping of meanings for unknown verbs from their argument structure
Phonological Production	Accurate production of phonetic features of English common to MAE and AAE

These aspects of English were selected because they represent core features of the grammar, semantics, and functional use of the language that are non-contrastive between AAE and MAE. They therefore provide a rich assessment of children's knowledge of the language that is not in any way biased against speakers of AAE. Furthermore, they undergo considerable developmental growth across the age range of 4 through 9 years and discriminate clearly between children who have language difficulties and those who do not. In the following sections of the paper we will briefly describe three of the subtests to illustrate both the novel testing procedures and the dialect-neutrality of the assessment.

The data that we will show comes from the standardisation sample collected by speech-language pathologists from all regions of the USA for the DELV-SC and the DELV-CR. A total of 1014 four to nine year old children were tested: 217 four-year-olds, 266 five-year-olds, 300 six-year-olds, 56 seven-year-olds, 101 eight-year-olds, and 74 nine-year-olds. Younger children were more heavily sampled because it is between the ages of 4 and 6 years that many of the features assessed by the DELV are acquired by typically developing children. It is also during this period of time that most children with possible language impairment come to the attention of teachers and clinicians and receive language assessments. Approximately 60% of the children were characterised by the testing clinicians and the schools as speakers of AAE, the other 40% as speakers of MAE. All of the AAE speakers were African American and most of the MAE speakers were white, though some of the MAE speakers came from Asian American and Hispanic backgrounds. The two dialects were evenly spread across the different age groups. AAE- and MAE-speaking children were matched for the education level of their parent(s) (typically some high school or completed high school). Approximately one-third of all the children at each age and for each dialect group had been identified by the participating clinicians as language-impaired and were receiving language intervention services.

2.2.1 Wh-question Comprehension

The grammar of wh-questions in English incorporates several fundamental aspects of syntax – the nature of syntactic variables, the distinction between adjuncts and arguments, rules governing the movement of syntactic categories in the formation of different sentence types, and the blocking of such movement by various elements in complex sentence forms (De

Villiers 1996). Wh-words function as syntactic variables in that they refer not just to single referents, but to the members of a set of referents. Consider the following example. You are told that Peter went to the clothing store and bought a shirt, a pair of shorts, and a belt. Later you are asked: "What did Peter buy at the store?" The correct answer is to list all the things he bought, not just one or some of them. This property of wh-words is particularly clear in double wh-questions. Told that Peter bought a shirt and a tie and Jill bought a dress at the clothing store, you are later asked: "Who bought what?" Now you have to give a "paired set" answer, pairing up all the members in each set in an ordered relation: "Peter bought a shirt and tie and Jill bought a dress".

According to modern generative syntax, wh-question formation involves movement of the wh-word from various places in the underlying structure of the sentence, leaving behind a trace of the missing sentence constituent to which it refers (Chomsky 1977). Short distance movement refers to the simple movement of a wh-word from the end to the beginning of a single clause, as in *Where_i did he go t_i?*. In long distance movement, the wh-word moves over two clauses to the front of the sentence: *Where_i did he say he went t_i?*. Some complex wh-questions are ambiguous as to whether the wh-word refers to the short or the long distance trace constituent. So a question such as *When_{i or j} did the man say t_i he was going to town t_j?* could be asking either when he said it or when he was going. Other contextual and discourse information would need to be used to disambiguate the question since the syntax allows both interpretations.

However, certain structures function as "barriers" that block the movement of wh-words out of clauses (Chomsky 1986). So *When did the man say **how** he was going to town?* is not ambiguous. It can only mean "When did he say it?" not "When was he going to town?". The second wh-word between the two clauses blocks the long distance interpretation since it serves as a syntactic barrier to the movement of the wh-word out of the lower clause. Various adjunct clauses such as relative clauses and purposive adverbials also function as barriers to the extraction and movement of wh-elements from clauses. De Villiers and Roeper (cf. De Villiers 1996; De Villiers and Roeper 1995; De Villiers, Roeper, and Vainikka 1990; Roeper and De Villiers 1994) have shown that young children are sensitive to these constraints on wh-movement at an early age.

The wh-question comprehension subtest of the DELV-CR assesses children's mastery of each of these syntactic features of questions in English: wh-words as variables, long distance movement, and sentence medial wh-words and relative clauses as barriers to wh-movement. On each item the child is told a brief story about a series of events shown in a sequence of two to four pictures. S/he is then asked a key test question about some aspect of the event. The pictured events and the stories describing them are designed to support several possible interpretations of the test question, so the child has to process the crucial syntax of the form in order to answer it correctly.

For example, to assess children's mastery of double wh-questions, they might be told the following story (accompanied by a picture of a girl eating a lollipop and a woman eating a chocolate bar)⁶: "The little girl and her mother were each eating some candy. Who ate what?". Younger children and language-impaired children tend to give singleton answers to the questions, e.g., "A lollipop" or "A chocolate bar", or they answer only the "who" or the "what" part of the double wh-question: "The girl and the mother" or "A lollipop and a chocolate bar". At other times, they provide only one pair, e.g., "The girl ate a lollipop". However, by age 6 almost all of the typically developing children correctly respond to the question with a "paired set" answer: "The girl ate a lollipop and her mother ate a chocolate bar".

The following example from De Villiers and Roeper (1996) illustrates the kind of item assessing understanding of barriers to long distance movement of the wh-word: "This boy loved to climb trees in the forest (picture of a boy climbing a tree). One afternoon, he slipped and fell to the ground (picture of the boy having fallen out of the tree). He picked himself up and went home. That night when he had a bath, he found a big bruise on his arm (picture of the boy in the bath with the father looking on holding a towel). He said to his dad, I must have hurt myself when I fell this afternoon." "When did the boy say how he hurt himself?" Since the medial "how" blocks any long distance interpretation of the when question, the only grammatically acceptable answer in English is "that night" or "in the night" or "when he took a bath", not "in the afternoon" or "when he fell out of the tree". Younger children and language-impaired children more often answer the question wrongly, and in particular, they

answer the middle "how" question. The latter response is an acceptable alternative in some languages, but not in English.

In the standardisation testing for the DELV-CR, the question comprehension subtest proved to be an effective dialect-neutral assessment of the children's knowledge of these core features of the syntax of English. As is shown in Figures 1 and 2 overleaf, this procedure produced all three features desired of such an assessment. There was strong developmental growth over the target age range, language-impaired children scored significantly lower at each age, and there was no difference in performance between the two dialect groups. Analysis of variance revealed significant main effects for age ($F(5,990)=45.1$, $p<.001$) and clinical status ($F(1,990)=51.5$, $p<.001$), but no effect for dialect ($F(1,990)=.01$, ns) and no significant interactions among the variables.

2.2.2 Wh-question Asking

Between the ages of 2 and 6 years, children master a variety of wh-questions that request different kinds of information from the listener – specification of objects ("what"), persons ("who"), locations ("where"), causes ("why"), times ("when"), or instruments and manners of action ("how") (De Villiers and De Villiers 1985; James and Seelbach 1982). Asking appropriate questions to obtain needed information is a vital skill for early school success. As part of the pragmatics assessment on the DELV project, we developed a question elicitation procedure to assess children's ability to ask the right question to find out missing information about a pictured event.

On each trial of the question asking test, the child is shown a picture with a crucial element missing from it. The missing piece of the picture is indicated by an ambiguously shaped blank white space surrounded by a dotted line. The "game" calls for the child to ask "the right question" to discover what is happening in the picture. In order to motivate the different types of wh-question, the missing elements of the pictures include objects, people, locations, tools, and causes of a character's emotion. At the beginning of the question-asking game, two warm-up items are used with a great deal of prompting by the tester to elicit "what" and "who" questions from the child. These items are designed to get the child to ask questions rather than guess at the answer (a predominant response from preschoolers). Nine test trials follow the

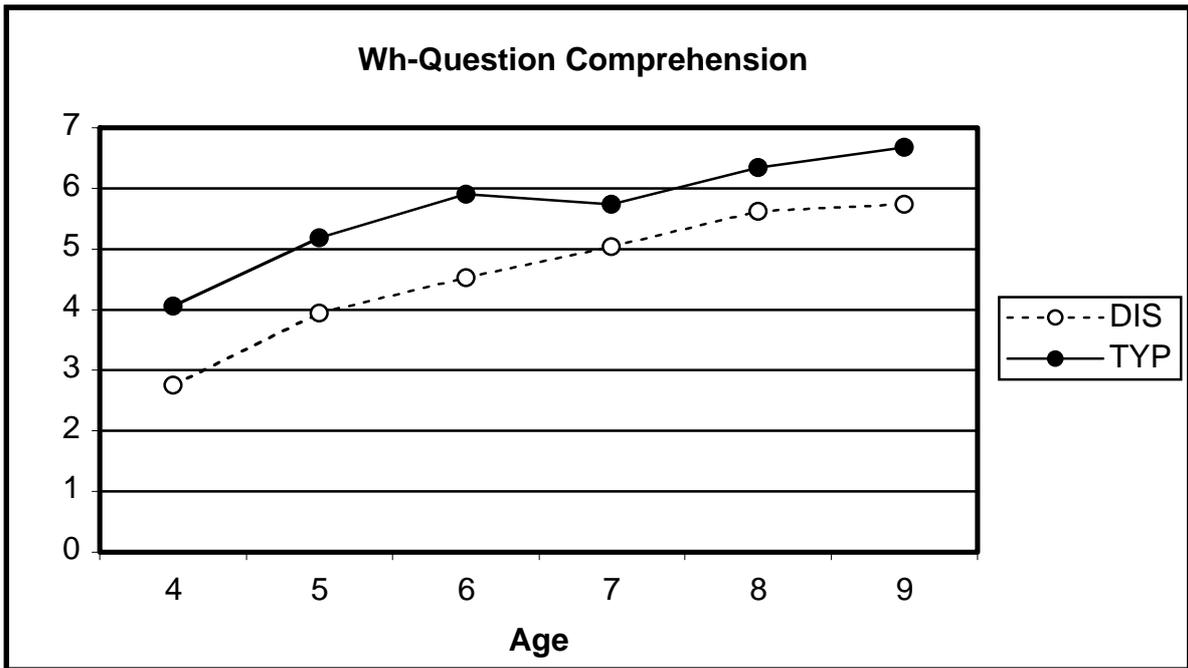


Figure 1. Performance of typically developing and language-impaired children on question comprehension

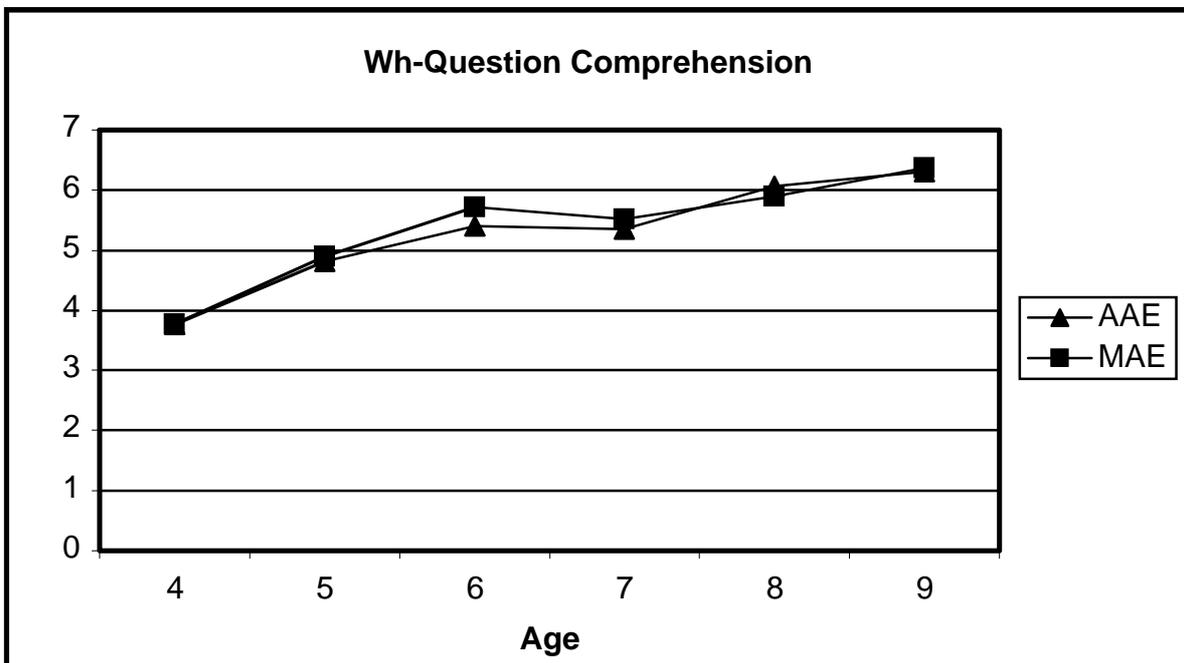


Figure 2. Performance of AAE- and MAE-speaking children on question comprehension

warm-up, with each picture designed to elicit a different wh-form. The last of these items provides pragmatic and referential motivation for the child to produce a double wh-question ("Who is eating what?"). The amount of prompting provided by the tester varies across the trials as the task progresses, from a great deal of verbal scaffolding to just the referential support provided by the picture (De Villiers, in press).

The children's responses on this task are scored for whether they produced a semantically and pragmatically appropriate question, not for the morphosyntax of the form that the child used. This allows for dialect variation in the morphosyntax of questions between AAE and MAE. So "What she paintin'?" (AAE) receives the same credit as "What is she painting?" (MAE). In this way, the pragmatics test is not biased against the AAE speakers. Empirical verification of this lack of bias is shown in Figure 3 overleaf. The task revealed substantial developmental growth across the age range, but the growth functions for the two dialect groups fall on top of one another. On the other hand, the question asking subtest clearly discriminated between the typically developing children and the language-impaired group at each of the ages (cf. Figure 4 overleaf). Statistical analyses indicated robust main effects for age and clinical status (Age, $F(5,990)=54.9$, $p<.001$; Clinical Status, $F(1,990)=84.7$, $p<.001$). However, there were no significant effects of dialect ($F(1,990)=.07$, ns), and none of the interactions between age, dialect, and clinical status were significant.

The typical errors of the youngest and most severely language-impaired children in the standardisation sample were failures to ask questions on many of the items, guessing at the answer, or simply saying "I don't know". A little more advanced developmentally were responses that asked a wh-question, but either an all-purpose one like "What is happening?" or "What is she doing?", or the wrong wh-question to get at the information needed. The most difficult item was the one designed to elicit a double wh-question.

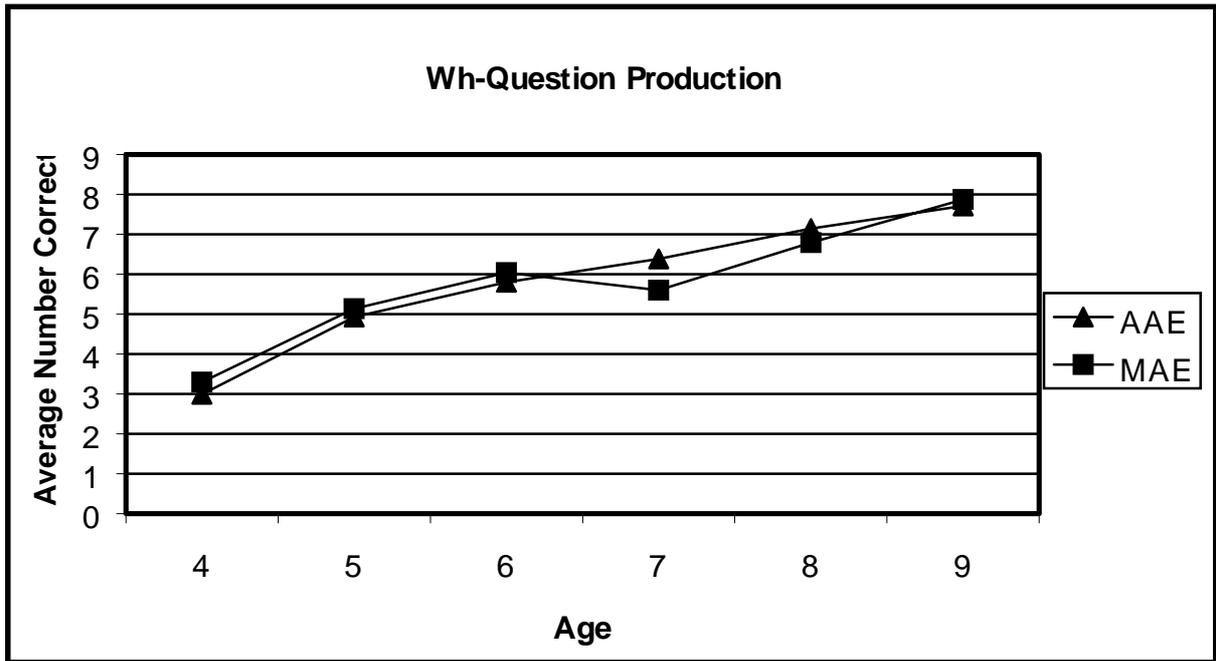


Figure 3. Performance by AAE- and MAE-speaking children on the question-asking task

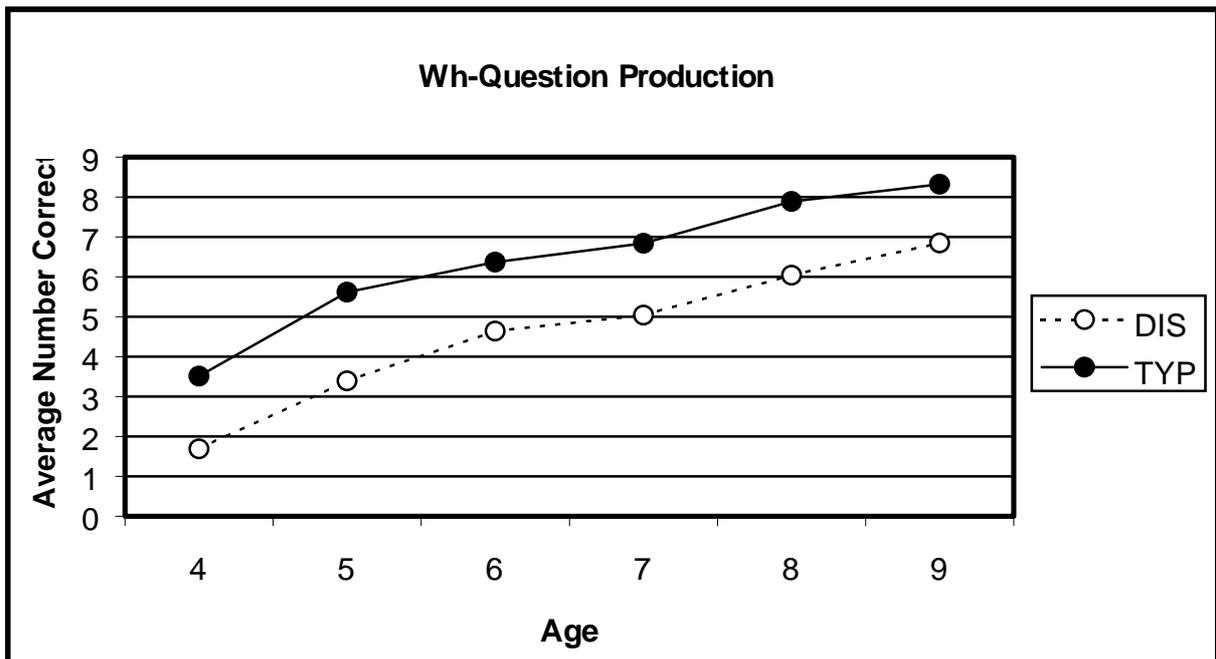


Figure 4. Performance of typically developing and language-impaired children on the question-asking task

2.2.3 Fast Mapping of Word Meaning by Syntactic Bootstrapping

We argued above that acquired vocabulary is too culturally dependent and so most vocabulary tests are biased against many dialect-speaking children. Thus, African American children frequently score lower than their white peers on widely used picture vocabulary tests (Stockman 2000). In the present project, we addressed this issue in several ways. Two of the semantic subtests focus not on number of words comprehended or produced in picture labeling situations, but on the organisation and relationships expressed between words in the children's lexicon. These tests use verbs and prepositions since it is likely that action and location words are more common aspects of these children's experience than noun domains such as animals, foodstuffs, or tools. Furthermore, especially for older children, lexical organisation may be as significant as size of vocabulary for identifying language problems (De Villiers, in press).

One of the most novel aspects of the DELV-CR is the subtest that assesses children's ability to "fast map" the meaning of a word from the linguistic context in which it appears. We reasoned that in order to determine whether a child has a language impairment with respect to vocabulary learning, it is more direct and culture-free to directly examine the process of learning words than to test which particular words have been acquired. It is now well-known that most of children's vocabulary after the first year or two of language acquisition is not acquired through parents or other people pointing out and labeling referents for the child. In fact, most words are learned from the conversational context in which they are embedded. Children must use the sentence structure as well as the referential situation and communicative context in order to determine what part of speech the word might be as well as some of its semantic properties. Normally developing children between the ages of 3 and 5 years seem to be able to do this from very few exposures to a word in context, hence the notion of "fast mapping" of at least partial word meanings (Carey 1978; Dollaghan 1985).

Linguistic context is particularly significant for the learning of verb meanings. Naigles (1990) demonstrated that 2-year-olds were sensitive to the differences between transitive and intransitive sentence frames in identifying the likely action labeled by a novel verb. Children hearing the sentence "Big Bird is gorging the bunny" looked more often at a videotape clip of the popular character from the television show Sesame Street performing a strange action on a

character dressed as a rabbit than at a clip of the two characters engaged in parallel independent actions, whereas children hearing "Big Bird and the bunny are gorging" looked more at the clip of parallel independent actions. Preschoolers can use the number of arguments in a sentence to determine whether a novel verb has a transitive or intransitive meaning, even when pronouns are substituted for the agent and object so that there is precious little lexical content present (Fisher 1996). Gleitman (1990) coined the term "syntactic bootstrapping" to characterise the process of narrowing down the likely meaning of a verb by the sentence structures in which it occurs. She argued that the general meanings of transitive or intransitive verbs, transfer (dative) verbs, and even verbs of communication and mental state (that take complement structures), could be determined by this process. Children with language impairment seem to have special difficulties in the process of fast mapping new meanings from linguistic context (Rice, Buhr, and Nemeth 1990), and they are prone to having particularly limited verb lexicons relative to their knowledge of nouns (Rice and Bode 1993).

The fast mapping test on the DELV-CR shows the children a sequence of three pictures of two characters performing unusual actions. The events are then described in a single sentence containing a novel (nonsense) verb. The stimulus sentences vary as to the type of argument structure provided. In intransitive contexts, the verb has a single argument ("The girl is temming"); in intransitive frames, there are two arguments ("The boy is temming the girl"); and in dative and complement frames, there are three arguments ("The mailman temmed the letter to the woman" and "The policemen temmed the man to stop the car"). Each picture sequence depicts at least two actions, sometimes involving action of one character on another and sometimes not. So, without the syntax of the stimulus sentence being taken into account, they are referentially ambiguous. Following the stimulus sentence, the children are asked a series of questions to see what meaning they attribute to the unknown verb. Depending on the particular stimulus sentence frame, these may include "Which one is temming?" (Agent), "Which one is the temmer?" (Agent), "Which one got temmed?" (Patient), or "Which one is temmable?" (Patient). The complement frames also have questions about the subject or the object of the complement: "Which one did he temm to stop the car?" and "Which one did he temm the man to stop?"

This is a difficult task for young children, so before the children see the items with strange actions and novel verbs, they are tested with real verbs and familiar actions. In this way, they learn the general idea of the task, and the testing clinician can determine whether any of the test questions pose particular difficulties for the children being tested even when they already know the meaning of the verbs being used in the stimulus sentences.

The data from the standardisation of the DELV-CR showed that the argument structures and the test questions varied in their difficulty for the children. Transitive/intransitive frames were easier than dative ones, and the latter were much easier than the complement structures, even though they include the same number of arguments of the verb. The test questions containing the derivational morpheme *-able* were harder than those with the other inflections. However, this allowed the task to work well as a semantic assessment for 8- and 9-year-olds as well as for younger children. There was steady developmental growth across the entire age range of 4 through 9 years, and no significant dialect differences were observed at any of the ages (cf. Figure 5 overleaf). However, as can be seen in Figure 6 overleaf, the task was clearly discriminating of language impairment. Analysis of variance revealed significant main effects of age ($F(5,990)=31.2, p<.001$) and clinical status ($F(1,990)=75.0, p<.001$), but there was no main effect of dialect ($F(1,990)=.35, ns$) and no significant interactions.

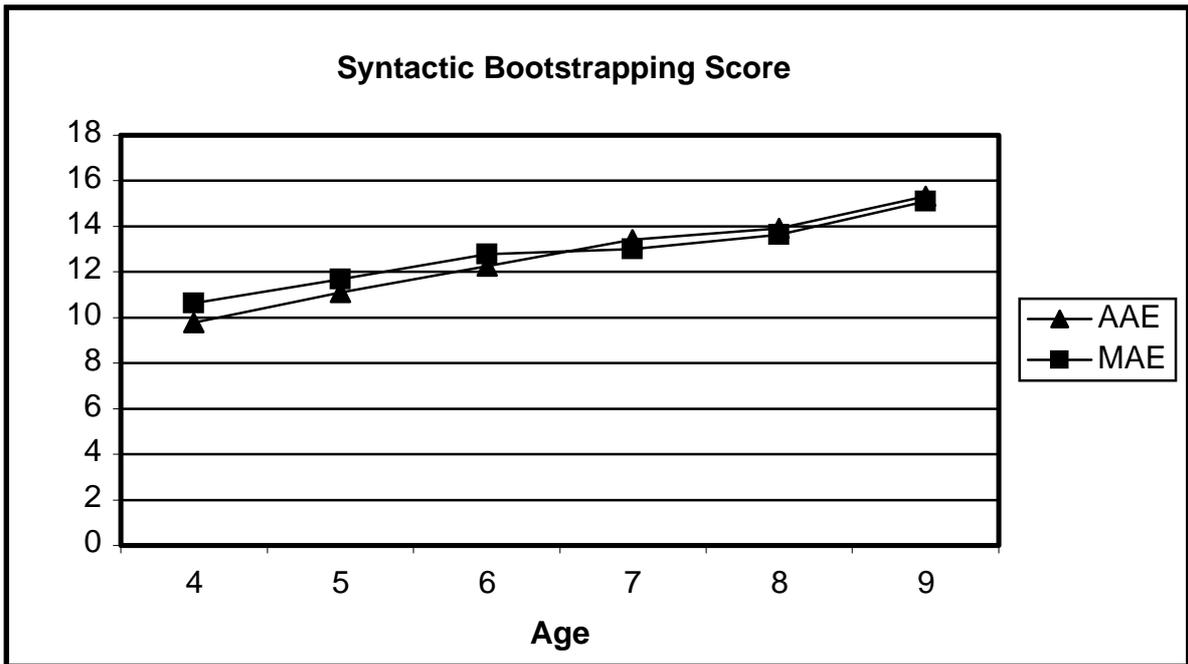


Figure 5. Performance of AAE- and MAE-speaking children on the fast mapping task

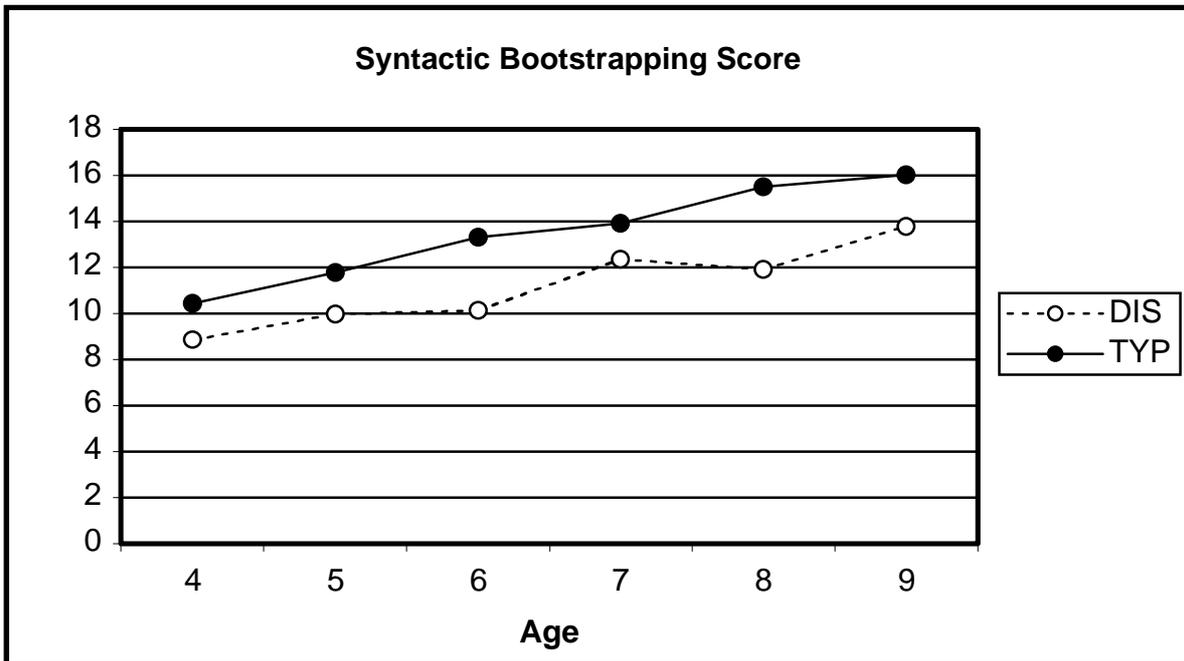


Figure 6. Performance of the typically developing and language-impaired children

3. CONCLUSION

We have demonstrated that it is possible to create a dialect-neutral assessment of first language acquisition in English by testing core features of syntax, semantics, and pragmatics that are non-contrastive across several dialectal variations of the language. The subtests of the DELV-CR each reveal strong developmental growth across the ages of 4 through 9 years, but there are no differences in the performance of AAE- and MAE-speaking children on any of the language tasks. Similar results are observed when the subtest data are aggregated into the language domains of syntax, semantics, and pragmatics.

The DELV-CR subtests provide the clinician with much more than just a diagnostic categorisation. They produce a substantial profile of the child's strengths and weaknesses in areas of the language that are important for success in early schooling and in the transition to literacy. As such, they supply a rich evaluation of language variation and its sources that has implications for areas and methods of language intervention as well as for assessment.

It remains to be seen how much these materials and testing procedures can be applied in their current form or with modification to other dialects of English besides MAE and AAE. However, we suggest that the principles and strategies adopted in this project are applicable for the development of new testing materials that will be far more appropriate and effective in multi-dialectal situations than the tests and procedures that are currently available.

NOTES

- ¹ See Green (2002), Washington and Craig (1994), and Wyatt (1995) for more detailed accounts of AAE and its acquisition by children.
- ² Some linguists refer to "Standard American English", but to avoid the prescriptive connotations of the notion that there is a standard form of the language, we prefer the term "Mainstream American English". MAE is simply one of several dialects of English; one which happens to have higher status than many other dialectal variations of American English through historical accident.
- ³ On these issues, see also Seymour, Bland-Stewart, and Green (1998); Taylor and Payne (1983); and Terrell and Terrell (1983).
- ⁴ This work was supported, in whole or in part, by NIH NIDCD Grant #R01 DC 02172-04 to Harry Seymour and Contract #N01 DC 8-2104 to Harry Seymour, Thomas Roeper, and Jill de Villiers, with The Psychological Corporation as a subcontractor.
- ⁵ We acknowledge the assistance of many consultants and graduate students on the project. They include Lamya Amdulkarim, Kristen Asplin, Linda Bland, Tim Bryant, Frances Burns, Tempi Champion, D'Jaris Coles, Mike Dickey, Christina Foreman, Debra Garrett, Lisa Green, Fred Hall, Bart Hollebrandse, Angelika Kratzer, Valerie Johnson, Janice Jackson, Caroline Jones, Minjoo Kim, Joe Pater, Eliane Ramos, Robin Schafer, Lisa Selkirk, Peggy Speas, Ida Stockman, Uri Strauss, Mike Terry, Shelley Velleman, Laura Wagner, and Toya Wyatt.
- ⁶ The example items given in this paper are similar but not identical to those on the DELV-SC and DELV-CR. Copyright agreements with The Psychological Corporation prevent us from using actual items from the tests.

REFERENCES

- American Speech-Language-Hearing Association. 1983. Position paper: Social dialects and implications of the position on social dialects. *ASHA* 25(9).
- American Speech-Language-Hearing Association. 2003. Technical report: American English dialects. *ASHA Supplement* 23.
- Campbell, T., C. Dollaghan, H. Needleman, and J. Janosky. 1997. Reducing bias in language assessment: Processing-dependent measures. *Journal of Speech and Hearing Research* 40: 519-525.
- Carey, S. 1978. The child as word learner. In M. Halle, J. Bresnan, and G. Miller (eds.) *Linguistic theory and psychological reality*. Cambridge, USA: MIT Press.
- Chomsky, N. 1977. On wh-movement. In P.W. Culicover, T. Wasow, and A. Akmajian (eds.) *Formal syntax*. New York: Academic Press.
- Chomsky, N. 1986. *Barriers*. Cambridge, USA: MIT Press.
- De Villiers, J.G. 1996. Defining the open and closed class program for acquisition: The case of wh-questions. In M.L. Rice (ed.) *Toward a genetics of language*. Mahwah: Erlbaum.
- De Villiers, J.G. In press. Cultural and linguistic fairness in the assessment of semantics. In H. Seymour and B.Z. Pearson (eds.) *Diagnostic evaluation of language variation: Distinguishing dialect and development from disorder. Seminars in Speech and Language*. New York: Thieme Medical Publishers.
- De Villiers, J.G. and T. Roeper. 1995. Relative clauses are barriers to wh-movement for young children. *Journal of Child Language* 22(2): 389-404.
- De Villiers, J.G. and T. Roeper. 1996. Questions after stories: Supplying context and removing it as a variable. In D. McDaniel, H. Cairns, and C. McKee (eds.) *Methods for assessing children's syntax*. Cambridge, USA: MIT Press.
- De Villiers, J.G. and P. de Villiers. 1985. Acquisition of English. In D.I. Slobin (ed.) *The crosslinguistic study of language acquisition: Vol 1: The data*. Hillsdale: Erlbaum.
- De Villiers, J.G., T. Roeper, and A. Vainikka. 1990. The acquisition of long distance rules. In L. Frazier and J.G. de Villiers (eds.) *Language processing and language acquisition*. Boston: Kluwer Academic Publishers.

- De Villiers, P.A. In press. Assessing pragmatic skills in elicited production. In H. Seymour and B.Z. Pearson (eds.) Diagnostic evaluation of language variation: Distinguishing dialect and development from disorder. *Seminars in Speech and Language*. New York: Thieme Medical Publishers.
- Dollaghan, C. 1985. Child meets word: "Fast mapping" in preschool children. *Journal of Speech and Hearing Research* 28: 449-454.
- Dollaghan, C. and T. Campbell. 1998. Non-word repetition and child language impairment. *Journal of Speech and Hearing Research* 41: 1136-1146.
- Fisher, C. 1996. Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology* 31: 41-81.
- Gleitman, L. 1990. The structural sources of verb meaning. *Language Acquisition* 1(1): 1-17.
- Green, L. 2002. *African American English: A linguistic introduction*. Cambridge: CUP.
- James, S. and M. Seelbach. 1982. The pragmatic functions of children's questions. *Journal of Speech and Hearing Research* 25: 2-11.
- Labov, W. 1972. *Language in the inner city: Studies in the Black English vernacular*. Philadelphia: University of Pennsylvania Press.
- Leonard, L.B. 1998. *Children with specific language impairment*. Cambridge, USA: MIT Press.
- Mesthrie, R. (ed.) 2002. *Language in South Africa*. Cambridge: CUP.
- Mufwene, S.S. 1998. *African-American English: Structure, history, and use*. London: Routledge.
- Naigles, L. 1990. Children use syntax to learn verb meanings. *Journal of Child Language* 17: 357-374.
- Rice, M.L. and J.V. Bode. 1993. GAPS in the verb lexicons of children with specific language impairment. *First Language* 13: 113-131.
- Rice, M.L., J.C. Buhr, and M. Nemeth. 1990. Fast mapping word-learning abilities of language delayed preschoolers. *Journal of Speech and Hearing Disorders* 55: 33-42.
- Rice, M.L. and K. Wexler. 1996. Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research* 39: 850-863.
- Rice, M.L. and K. Wexler. 2001. *Rice/Wexler Test of Early Grammatical Impairment*. San Antonio: The Psychological Corporation.

- Rickford, J.R. 1999. *African American vernacular English: Features, evolution, educational implications*. Malden: Blackwell.
- Roeper, T. and J.G. de Villiers. 1994. Lexical links in the wh-chain. In B. Lust, G. Hermon, and J. Kornfilt (eds.) *Syntactic theory and first language acquisition: Crosslinguistic perspectives. Vol. II: Binding, dependencies, and learnability*. Hillsdale: Erlbaum.
- Seymour, H., L. Bland-Stewart, and L. Green. 1998. Difference versus deficit in child African American English. *Language, Speech, and Hearing Services in Schools* 29(2): 96-108.
- Seymour, H. and B.Z. Pearson (eds.) 2004. Evaluating language variation: Distinguishing dialect and development from disorder. *Seminars in Speech and Language*. New York: Thieme Medical Publishers.
- Seymour, H. and C. Seymour. 1977. A therapeutic model for communication disorders among Black English speaking children. *Journal of Speech and Hearing Disorders* 42: 247-256.
- Stockman, I.J. 2000. The new Peabody Picture Vocabulary Test-III: An illusion of unbiased assessment? *Language, Speech, and Hearing Services in Schools* 31(4): 340-353.
- Taylor, O. and K. Payne. 1983. Culturally valid testing: A proactive approach. *Topics in Language Disorders* 3: 8-20.
- Terrell, S. and F. Terrell. 1983. Distinguishing linguistic difference from disorders. *Topics in Language Disorders* 3: 1-7.
- Van der Lely, H. 2003. Do heterogeneous deficits require heterogeneous theories? SLI subgroups and the RDDR hypothesis. In Y. Levy and J. Schaeffer (eds.) *Language competence across populations: Toward a definition of specific language impairment*. Mahwah: Erlbaum.
- Washington, J. and H. Craig. 1994. Dialectal forms during discourse of poor, urban, African American preschoolers. *Journal of Speech and Hearing Research* 37: 816-823.
- Wolfram, W. 1991. *Dialects and American English*. Englewood Cliffs: Prentice Hall.
- Wyatt, T.A. 1995. Language development in African American English child speech. *Linguistics and Education* 7(1): 7-22.
- Wyatt, T.A. 2002. Assessing the communicative abilities of clients from diverse cultural and language backgrounds. In D.E. Battle (ed.) *Communication disorders in multicultural populations*. Third edition. Boston: Butterworth-Heinemann.

