NIH Contract N01-DC8-2104
4/12/02

Developmental Milestones for AAE speaking 4-, 5-, and 6-year-olds.

University of Massachusetts Working Groups on African American English
Amherst MA
Harry N. Seymour, PI
Thomas Roeper (*UMass*) and Jill G. de Villiers (*Smith College*), Co-Collaborators
Peter de Villiers (*Smith College*), Consultant
Barbara Z. Pearson, Project Manager

in conjunction with
The Psychological Corporation
San Antonio TX
Lois Ciolli, Research Team Project Director

Consultants and Contributors:
Lamya Abdulkarim, UMass Amherst
Brenda Allen, Smith College
Kristen Asplin, UMass Amherst
Elena Benedicto, Purdue University
Timothy Bryant, UMass Amherst
Frances Burns, UMass Amherst
Robert Carlson, TPC, San Antonio TX
Tempii Champion, Univ. S. Florida
D'Jaris Coles, Wayne State University
Mike Dickey, Northwestern University
Juanita Doty, Jackson MS
Christina Foreman, UMass Amherst
Debra Garrett, College of NJ
Lisa Green, Univ. of Texas, Austin
Fred Hall, Worcester State College
Bart Hollebrandse, Univ. of Utrecht
Janice Jackson, Univ. of S. Carolina
Valerie Johnson, Vanderbilt University
Caroline Jones, Sydney, Australia
Minjoo Kim, UMass Amherst
Angelika Kratzer, UMass Amherst
Gay Lamey, TPC, San Antonio, TX
Glen Larson, TPC, San Antonio, TX
Matt Morris, TPC, San Antonio TX
Deanna Moore, UMass Amherst
Gail Nickel, TPC, San Antonio TX
Joe Pater, UMass Amherst
Eliane Ramos, Florida International Univ.
Robin Schafer, Yale University
Lisa Selkirk, UMass Amherst
Peggy Speas, UMass Amherst
Ida Stockman, Mich. State Univ.
Uri Strauss, UMass Amherst
Linda Bland-Stewart, GWashington Univ.
J. Michael Terry, UMass Amherst
Shelley Velleman, UMass Amherst
Laura Wagner, Harvard University
Charles Wilkins, TPC, San Antonio TX
Toya Wyatt, Cal State, Fullerton

Editorial Assistants, UMass Dept. of Communication Disorders
Aimee Diehl
Sonia Gonzalez
Kate MacDonald
Lindsay Maxon
Nicole Mosely
Nicole Patton
Anne Ryder
Melissa Torres

Programming and Data Management
Paul Speckels

OBJECTIVES
(from the Technical Proposal, pp. 61-62)

Objective II.d **Analyze data, in particular regarding age-related acquisition of specific language milestones which are able to differentiate among age groups.**

(1) the statistical and psychometric analyses,
(2) follow-up analyses,
(3) interpretation, synthesis, and application.

Objective II.e. **Provide summary of development of phonology, syntax, semantics and pragmatics in child AAE**

Background:

The set of language behaviors to be used as the basis for the assessment instrument being developed by this contract was presented in the 4th semi-annual report of March 2000. Since that time, a preliminary testing of the proposed items was carried out on a group of approximately 800 4-, 5-, and 6-year-olds, 100 more than specified by the contract. Of that number, 400 are typically developing African American children, whose performance by 6-month age groupings on the 320 items of the tryout test provides the core data for this report. Their data are referenced against 3 other groups, a group of 130 language impaired AAE children, and two groups of SAE-speaking children matched for socio-economic class, age, and language status. In addition, their age trends are extended, where relevant, by similar analyses of data from 500 7- to 12-year-olds collected by our subcontractor, the Psychological Corporation, using the same methods and materials.

First, we document how the notion of a useful "language milestone" was operationalized to establish a selection of them for use at age 4, others at 5, and still others at 6. In the process, we have identified language behaviors that are 1) dialect specific and others that are 2) "dialect neutral"; that is, target behaviors that are SPECIFIC TO AAE and those that are"DEEPER THAN DIALECT." The first set of behaviors is useful for identifying AAE speakers and the course of development of AAE patterns. The second set is useful for *confirming typical language development* among 4-, 5-, and 6-year-old children learning several different varieties of English. In Phase 3 of the contract, we will refine the analysis of those items to select the ones capable of reliably differentiating delayed development. The first essential step in that process, though, as prescribed by the RFP, was to establish the set of typical behaviors for AAE children learning language "normally" (Phase I). The second, reported now at the end of Phase II, was to bracket the developmental ranges of those behaviors.

Please note that this is far from the definitive volume on AAE child speech, but it achieves the clear statement of a large number of AAE child speech features calibrated by age of mastery. The candidate behaviors are consistent with developmental trends identified by theoretical psycholinguistics and traditional clinical practice, the two disciplines from which they derive. They can be now be implemented (in Phase III) as diagnostic markers for language impairment in the African American speech community, and perhaps in other dialect communities as well.

Overall Results:

The goal of this project has been to create a test which allows non-contrastive features of grammar to emerge as decisive in the evaluation of children. There are two kinds of non-contrastive features: direct representation of linguistics universals, and features which are the same in all dialects of English. Eventually we hope to capture them both. The diagnostic elements of the DSLT focus on features of English common to both AAE and SAE, and therefore they allow us to capture just those features which clearly mark disorders.

The broad results are stunning. They show that typically-developing AAE and SAE speakers perform the same with respect to sophisticated syntactic and semantic aspects of their grammars. It is however notable that the Disordered AAE speakers---as identified by existing instruments--show better knowledge than the SAE disordered group.

Why should that be the case? The answer may lie in the difficulties of the previous methods of evaluation. Existing tests focus primarily on morphology which is precisely where dialect differences can be found. (See the discussion of ID features at Subtest 1.) Because they have been identified by morphology, some children appear to be disordered who are in fact not disordered. They can show the extent of this misdiagnosis when they have the opportunity to perform tasks that call upon their deep syntactic and semantic knowledge. The skills tested by the DSLT and reported here involve the most sophisticated aspects of grammatical knowledge that have ever, to our knowledge, been brought to a broad-based instrument used with school-age children.

The pattern of results shown in the graphs and statistics that follow is consistent with the hypotheses set forth in the Technical Proposal for this contract. Specifically, the preliminary findings confirm the following propositions:

1. Typically-Developing (TD) AAE and SAE children differ reliably in their production of contrastive grammatical morphemes.
2. The numbers of contrastive morphemes produced by TD-AAE and Language Disordered (LI) SAE children do not differ reliably at all ages (and they did not differ at all at age 4).
3. The same TD-AAE children whose pattern of morpheme use is similar to LI-SAE children will be observed to pattern WITH THE TD-SAE children (and NOT with the LI children) on a large number of abstract linguistic variables tested in the various subtests of the DSLT.

A non-exhaustive list of examples of such "abstract" linguistic variables (with the subtests where they are illustrated in parentheses) is as follows:

a. The use of "was/were" in copula and auxiliary contexts (MS1):
b. Correct production of paired exhaustive answers to double WH-questions (WH2);
c. Respecting barriers to long-distance movement in WH-questions, especially in sentences involving movement over WH-complements (WH2);
d. Understanding simple and complex properties of passive sentences (object as subject, hidden agents, disjoint reference, ongoing versus result readings) (PA3);

e. Decline of active-for-passive answers in comprehension of passive sentences (PA3);
f. Use of definite articles when required in "familiar-the" and "part-the" contexts (AR4);
g. Accurate and specific reports of speech acts, telling, asking, and prohibiting (CR5);
h. Understanding and reporting Theory of Mind of characters in a narrative; explaining actions by reference to mental events (SN6);
i. Growing use of sequencers (and adverbial clauses) to explain temporal sequence in reported events (SN6);
j. Adequate specification of referents involving locations of individuals (using prepositional phrases or relative clauses) on first pass answers, and with additional prompting (RS7);
k. Producing appropriate and specific WH-questions to obtain relevant information about a pictured context when prompted by wh-words, semantic domain clues, and no specific prompt both on first-pass answers and with additional prompting (QA8);
l. Producing two or more verbs (vocabulary) at varied levels of specificity relative to the specificity of a verbal prompt in relation to a single picture (VC9);
m. Providing an appropriate lexical or grammatical contrast among expressions involving prepositions (in response to different verbal prompts) (PC10);
n. Confining the scope of the word "every" to within a single sentence (and not across sentence boundaries) (QN11);
o. Understanding argument roles in real verbs of different argument structures (FM12);
p. Fast-mapping meanings of novel verbs according to the sentence frames they are used in (FM12);
q. Accurate repetition of non-words of 2 and 3 syllables (NW14).

Patterns of Interactions Observed:

As indicated in the graphs that make up the core of this report, some of the variables mentioned above show distinct developmental patterns for TD versus LI children with no interaction of group (ie. all groups show the same slope at different levels at all ages). Other variables exhibit interactions generally consistent with the diagnostic perspective of the contract. That is, the TD children may show consistent development across the age range, while the LI children do not; OR the TD and LI groups may all show development but the slope and/or timing of the observed growth may differ.

Limitations:

Subsequent analyses, principally by subcontractor The Psychological Corporation will evaluate which patterns of difference are reliable enough for diagnostic purposes in the context of the Phase 3 Data Collection. Other variables from the subtests, not listed above, may prove useful

for diagnosis at ages above 6, or they may provide valuable information for therapy goals. Subsequent reports will explain the meanings of the variables and the interpretation of the developmental patterns for a broader audience.

The goal of the report at this stage is to establish for the Working Groups themselves which variables should be carried into Phase 3 and which ones will need further development before they can be useful for diagnosis (and to report that much to our funding source). We recognize that many of the labels will appear opaque to a general reader and that the meaning of the graphs will not be as evident even to our close colleagues as they are to us. We welcome specific questions to help us improve the readability of this report in subsequent iterations.

Caveat:

THIS MANUSCRIPT IS ONE STEP ABOVE AN INTERNAL DOCUMENT. IT IS A WORK-IN-PROGRESS, NOT A PUBLICATION. The exercise of putting everything in a readable format has been helpful in showing the loud, clear message even to ourselves. But in the scheme of things, it is premature. It has not been subjected to the many checks and balances necessary for publication. The data were collected at considerable effort over a period of a year. They were another six months being entered and checked for consistency and so forth. This report represents only the 10th or 20th pass through the data. There will be many more passes through. The checking that happens to ensure accuracy in publication will happen in the natural course of the next steps, which we are already engaged in.

We have made every effort (and more) to be as accurate as possible. But we expect that many findings reported here will be amended as we continue to look at the data from different perspectives and become more aware of its many subtleties. The "level of polish" of the report is consistent with our urgent need to move to the next step: to use these data to refine the instrument that will go back out in the field in the fall. As we mention above, there is a "language sample" of sorts averaging 80-100 utterances for each child. There are 20 pages of variables created from them already. There is much more to be learned and said about them.

Caveat 2:

H. Seymour, Chair
Dept. of Communication Disorders
University of Massachusetts
Amherst MA 01003 USA
413.545.0131
fax: 413.545.0803
hseymour@comdis.umass.edu

or
Barbara Zurer Pearson, Project Manager
413.545.5023
fax: 413.545.0803
bpearson@comdis.umass.edu

The organization of this document is as follows:

I. Instrument
II. Subtests
III. Subjects (group labels)
IV. The Database
V. The Analyses
VI. The Report

I. The Instrument

The Tryout Version of the DSLT, comprised of 14 subtests and 320 items, was administered in a standardized protocol to 830 children nationwide. Their responses provide the basis for the analyses to follow. The test was professionally edited, and except for its size, was similar in format (but not content) to other language tests currently in the public domain: that is, it consisted of a stimulus book and a record book to record the answers by hand. No special technology was required for its administration. One hundred of the examiners were asked to provide audiotapes of some of the subtests, and an additional twenty in the Northeast were videorecorded. The stimulus book, record form, and administration directions accompanied Semiannual Report 6 (March 2001)

II. The Subtests

The four main linguistic domain divisions are represented in the subtests as follows:

Syntax: Subtest 2, WH Syntax (WH)
Subtest 3, Passive (PA)
Subtest 4, Articles (AR)

Semantics: Subtest 9, Verb Contrasts (VC)
Subtest 10, Preposition Contrasts (PC)
Subtest 11, Quantifiers (QN)
Subtest 12, Fast Mapping (FM)

Morphosyntax and Phonology:
Subtest 1, Morphosyntax (MS)
Subtest 13, Phonology (PH)
Subtest 14, Novel-word Repetition (NW)

Pragmatics: Subtest 5, Communicative Role Taking (CR)
Subtest 6, Short Narratives (SN)
Subtest 7, Reference Specification (RS)
Subtest 8, Question Asking (QA)

III. Subjects

The Phase 2 data collection targeted 690 children of 3 distinct language backgrounds:

- typically developing AAE-speaking children (TD-AAE);
- typically developing Standard American English speakers (TD-SAE); and

- AAE-speaking children identified in their communities by speech pathologists, teachers, and parents as language impaired (LI-AAE).

The actual "UMass" sample includes 830 children: 631 of the targeted 690: that is, 44 fewer TD-AAE children, 68 extra LI-AAE children, and a new category of 86 LI-SAE learning children. There are also 25 TD-AAE children who participated in the Inter-rater study and took the test twice. Their second administration is not counted in these figures. The tables, with the 6-month age breakdowns and information about the Parent Education Levels (PED), ethnicity, region, and gender of these children are included in the Subcontractor's Report and the Annual Technical Report (enclosed with this report).

I include the Developmental Milestone Study numbers by language group and age group here for easy reference.

Table 1.

"Study"	4;0-4;5	4;6-4;11	5;0-5;5	5;6-5;11	6;0-6;5	6;6-6;11	Total
TD-AAE	64	61	65	71	75	70	406
TD-SAE	32	33	29	38	47	30	209
LI-AAE	10	20	25	24	28	22	129
LI-SAE	9	12	14	17	12	22	86
Total	115	126	133	150	162	144	830

These groupings have permitted us to do the four-way analysis of the selected variables that appear in the body of this report.

In addition, there are 700 additional children tested by The Psychological Corporation with the same materials. Of those "TPC subjects," we use 486 who are categorized in the language groups above as a reference in the graphs below. They are the ones included as "7 to 12" where they appear. They have been especially useful in the analysis of the trends for the 6-year-olds of the

Note about the (inevitable) confusion of the labels:

Where possible we have adopted the labels outlined above, TD versus LI (typically-developing versus language-impaired) which have good currency in the literatures of Communication Disorders and Child Language. However, many of the materials from which we worked had the labels used by The Psychological Corporation (TPC): "Try" (for "Tryout Study), which means "typically-developing" and "Dis" for "disordered." Since we continue to go back and forth with TPC with these data on a daily basis, in order to facilitate communication and maintain consistency as we move from one iteration of the data to the next, we have not transformed the labels in their files. Therefore, the graphs, made directly from those files tend to have the TPC labels; our discussion has the labels of the field. The reader is asked to consider them as synonyms.

IV. The Database

The "UMass Dataset" consists of the responses and associated coding for all questions for UMass subjects and the supplemental TPC subjects, minus subjects with incomplete data or who were tested but then disqualified for other reasons. Most of those taken out are retained in a

"flagged" file where they can be analyzed independently later, if it is relevant. The UMass Dataset is originally stored in SAS by TPC mainframe programmers and transferred to UMass in intermediate Excel files by ftp. UMass assembled them in Filemaker Pro 5.0 (on a MAC G3), and then re-output them to the authors in Excel files, which they could then analyze in Excel or import into various statistical program on their personal computers, some MACs and some IBMs.

The original SAS file contains over 6000 variables for each of 1700 subjects. Of those, approximately 2500 variables are relevant for the UMass authors. TPC programmers split the file in groupings by subtest, with a 15th grouping for "subjects." (Note that three of the subtests' response files exceed the 65,000 maximum for records per file permitted by Excel). Responses to all subtests except 3 were in the form of verbatim transcripts of the child's response, which varied from 1 word to as much as 2 sentences, or over 80 such utterances per child. The Filemaker database is organized by subtest, plus a SubjectMaster file. Each subtest has an "ItemsN" file, with information about the categorization of the items; a "ResponsesN" file, with each response for each subject and its associated coding, most of it done by hand, some of it accomplished through logic scripting; and finally, a "SummariesN" file, with summary data, various totals, subtotals, and tallies of specific codes for each subject. In the Filemaker database, there is also the skeleton of a "GrammarMaster" section for summarizing data by principles that span across subtests. Thus far, only a few cross-subtest analyses have been done, using the provision in Filemaker to create relationships across files and then exporting to a common file. The GrammarMaster facility will be more developed in the next steps of the data analysis.

I attach a copy of the DSLT list of variable names by subtest (Appendix 1) and the Coding Principles (Appendix 2).

V. The Analyses

To determine which developmental trends were sharp enough to be "milestones," we began by graphing each composite variable (and some individual items) by group and by age. Although the focus was on the performance of the 400 typically developing AAE children, we graphed them against the background of the typically developing SAE speakers and the two disordered groups. The original graphs went from age 4 to 12 in year intervals. From the preliminary graphs, we selected the variables that were consistent with our hypotheses about the task and showed a coherent pattern across the 4 to 6 age range (either rising, falling, or level). We regraphed them in 6-month intervals during the target age range, retaining the older children as a single averaged point to show the direction of development after 6;6.

To get a better sense of what patterns were meaningful (and reliable), we selected a few of the major variables in each subtest and compared the means by age and group in a multivariate analysis of variance (SPSS 9.0 for Windows). This program produced a significance level for the overall comparison as well as for pairwise comparisons down to each 6-month interval within group. ALTHOUGH SOME GRAPHS INCLUDE THE 500 SUBJECTS OVER AGE 7, THE STATISTICS WERE DONE ONLY ON THE CHILDREN IN THE AGE RANGE 4-6.
These p-values were especially helpful, given the different sample sizes within the groups. Naturally, smaller differences could be reliable for the TD-AAE sample, which included approximately 70 subjects per cell, than for the LI-SAE sample with only 15 or so at each point. The statistical analysis helped us recognize what magnitude of mean difference on over a hundred

different "scales" was potentially meaningful and which mean differences were most likely products of random variation.

For the Identifier items in the Morphosyntax and Phonology subtests, we focused on the two TD groups; for the Diagnostic items (most of the test), we focused on the patterns of development in the two AAE groups, TD and LI. The LI group gave us a sense of what differences could be discernible in the lower tail of the TD developmental distribution. However, since the diagnosis of LI in the AAE population is problematic (!), we also included the comparison to the LI-SAE group, whose diagnosis is more reliable. This turned out to be especially helpful in the 4- to 5-year-old age range, where children presumably have fewer other external or academic signs of language deficits and so for the AAE children the lack of knowledge about typical developmental patterns of the dialect community is more likely to OVER-identify LI children. In fact, as one will readily see, the data for many of the variables reported here are consistent with this speculation.

VI. The Report (Pages 1-0 to 14-5).

Information included:

The subtest is the unit of organization for the report. Each subtest has a cover sheet reminding the reader of where the subtest fits into the whole, the basic format of the items, the major coding principles (details are found in Appendix 2), the list of figures, and a general discussion of the findings to orient the reader. The major variables are presented one to a page with a group by age line chart depicting the 4 major groups of interest (and including the 7- to 12-year-olds as a single averaged point, just for reference) and a "detail" chart (figure "b") focused more clearly on the Typically-Developing AAE children. PRELIMINARY analyses are found below the charts: 1) F-statistics, done with all 4 groups over the six 6-month age levels, followed by 2) the mean difference and significance levels from the pairwise comparisons of the 4 groups. A third statistic reports a univariate test of the age effect when group is held constant.

The general format of the charts:

Since developmental milestones are the focus of the report, AGE is portrayed on the x-axis. The y-axis is the SCORE with the particular unit of measure for that chart "labeled" in the title, not directly on the chart. Most are "average number correct" in the score for whichever items are included in the subtotal designated by the title; occasionally there is a "percent correct," or a tally of a specific response code for the subtest (eg. "bare singulars" for the Article subtest or "singleton" answers in the WH-syntax). Where possible, the codes are explained on the page where they are shown, but a fuller explanation is available in the book of the Coding Guidelines which accompany this report.

The "detail" chart is generally a bar graph with the same axes as the accompanying line graph. Its function is to highlight the developmental pattern for the Typically-Developing (TD) AAE-learning children. For Identifier Items (in MS and PH), they are shown with the TD-SAE-learning (Standard American English) children. For the Diagnostic Items (all of the ot hers), the "TD-AAEs" are shown with the Language-Impaired (LI) AAEs.