NIH Public Access
Author Manuscript
Lang Speech Hear Serv Sch. Author manuscript; available in PMC 2012 July 16.

Published in final edited form as:
Lang Speech Hear Serv Sch. 2010 July ; 41(3): 328–339. doi:10.1044/0161-1461(2009/08-0077).

Index of Productive Syntax for Children Who Speak African American English

Janna B. Oetting,
Louisiana State University, Baton Rouge

Brandi L. Newkirk,
Louisiana State University, Baton Rouge

Lekeitha R. Hartfield,
Louisiana State University, Baton Rouge

Christy G. Wynn,
Louisiana State University, Baton Rouge

Sonja L. Pruitt, and
San Diego State University, CA

April W. Garrity
Armstrong Atlantic State University, Savannah, GA

Abstract

Purpose—The validity of the Index of Productive Syntax (IPSyn; Scarborough, 1990) for children who speak African American English (AAE) was evaluated by conducting an item analysis and a comparison of the children's scores as a function of their maternal education level, nonmainstream dialect density, age, and clinical status.

Method—The data were language samples from 62 children; 52 of the children were between the ages of 4 and 6 years and were classified as developing typically, and 10 were 6 years old with specific language impairment (SLI).

Results—All IPSyn items were produced by at least 1 child, and 88% of the items were produced by 50% or more of the children. The children's IPSyn scores were unrelated to maternal education level and dialect density and were visually comparable to IPSyn scores reported for children who speak mainstream English. Nevertheless, IPSyn could not be used to detect differences between the 4- to 6-year-olds based on age, nor could it be used to detect differences between the 6-year-olds with and without SLI.

Conclusion—IPSyn is a valid measure for AAE speakers, but it can be insensitive to age and clinical differences between children who are over the age of 48 months.

Keywords

AAE; assessment; child language impairment

© American Speech-Language-Hearing Association

Contact author: Janna B. Oetting, Department of Communication Sciences & Disorders, 64 Hatcher Hall, Louisiana State University, Baton Rouge, LA 70803. cdjanna@lsu.edu.

These data were presented at the 2008 annual convention of the American Speech-Language-Hearing Association in Chicago, IL.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Language sampling is an important component within assessment because conversing with a child is one of the best ways to learn how a child uses language to communicate with others. The truthfulness of this claim, however, is dependent on the validity of the measures generated from the language samples (Stockman, 1996). Although a number of language sample measures exist within the field of speech-language pathology, many of them have not been validated for children who speak one of the many nonmainstream dialects of English that are spoken in the United States and elsewhere. Moreover, of those measures that have been evaluated for their across-dialect validity, the results have not always been positive.

As an example, consider Lee's (1974) Developmental Sentence Scoring (DSS) system. The DSS system allows a clinician to index a child's grammatical development using a 50-utterance language sample. With this measure, children earn points based on the grammatical complexity of their utterances. Unfortunately, the scoring system that is used to calculate a child's DSS score is based on mainstream English. Given this (and perhaps not surprisingly), multiple studies have shown the DSS system to be invalid for assessing the grammatical development of children who speak nonmainstream English (Nelson & Hyter, 1990; Oetting, 2005; Oetting, Cantrell, & Horohov, 1999).

Early studies of a more recently developed measure, the Index of Productive Syntax (IPSyn; Scarborough, 1990), have shown some success at assessing the grammatical development of children who speak nonmainstream English. IPSyn is similar to DSS in that it allows a clinician to index a child's grammatical development from a language sample. Unlike with the DSS system, however, individual utterances are not scored for their grammatical components; instead, scoring is based on the child's use of 56 different types of grammatical structures, and these structures can appear anywhere within the child's analyzed sample. Another unique feature of the IPSyn system is that the child only needs to produce two different exemplars of each structure to receive maximum credit. Given this, children whose dialects allow for some grammatical structures to be optionally omitted or zero marked are not penalized if they produce these structures at least some of the time.

To illustrate why this feature of the IPSyn scoring system appears ideally suited for nonmainstream English speakers, consider the ways in which a child's grammatical development is indexed with the IPSyn and DSS systems using the following three utterances.

I'm good.

Him and Tyler mean.

They was in trouble all last week.

All three of these utterances are felicitous within a number of non-mainstream dialects of English. Nevertheless, only the first utterance contains a mainstream English exemplar of copular BE. In the second, the copular BE is zero marked, and in the third, the copular BE form, *was,* is used instead of the mainstream form *were*. With the IPSyn scoring system, any child who produced these three utterances within a sample would receive maximum credit for the copular BE structure (because the utterances contain two substantially different exemplars of copular BE). In contrast, each of these utterances would receive a numerical score within the DSS system, and scoring of these utterances would be based on the child's use of mainstream English structures only. Given this, the DSS total score of a nonmainstream English-speaking child who produced these three utterances would be lower than the DSS total score of a mainstream English-speaking child (assuming that the mainstream English-speaking child produced *am* in the first utterance, *are* in the second, and *were* in the third).

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Of course, just because IPSyn appears ideally suited for a group of children does not provide evidence of its validity. Therefore, there is a need to rigorously evaluate the validity of the measure using data from nonmainstream English-speaking children. In the current study, we did this using a new set of language samples from 62 children who spoke African American English (AAE). AAE is a model system for examining the effects of nonmainstream English on various child language measures because of its high rates of non-mainstream features relative to other nonmainstream dialects and the wide range of variability (i.e., low vs. medium vs. high) that exists among AAE speakers in their rates of nonmainstream feature use (Oetting & McDonald, 2002; Washington & Craig, 1994). As background, we further describe the IPSyn scoring system and the existing literature on this tool.

IPSyn as an Assessment Tool

IPSyn was designed to be used with children who are between the ages of 24 and 48 months. As mentioned earlier, it involves the scoring of 56 grammatical structures, and scoring is based on a child's use of these structures during a 100-utterance language sample. Each target structure can receive up to 2 points, and points earned can be combined for a total score and four subscale scores (i.e., noun phrase, verb phrase, question/negation, and sentence structure). Given that the scoring system is limited to no more than two sufficiently different tokens of each structure, IPSyn is typically described as a measure of a child's emergence of grammar rather than mastery.

The initial IPSyn study by Scarborough (1990) examined the measure's ability to detect age-related changes in children's development of grammar. The participants were 15 children who were learning mainstream English, and their IPSyn scores were calculated when they were 24, 30, 36, 42, and 48 months of age. As expected, the children's IPSyn totals increased as they aged. Age increases also were observed for all four subscales up until the children were 36 months. After this age, the children's scores on the noun phrase subscale did not increase significantly, and after 42 months, a similar plateau was observed for the children's scores on the verb phrase subscale. In other words, after 42 months, significant increases in the children's IPSyn scores were limited to the question/negation and sentence structure subscales.

Since Scarborough's study, IPSyn has been used to investigate the grammatical development of various groups of children, including typically developing toddlers (Horton-Ikard, Ellis Weismer, & Edwards, 2005; Rispoli & Hadley, 2001); children who are at risk for language delays (Thal, Reilly, Seibert, Jeffries, & Fenson, 2004); children who are late to talk (Rescorla, Dahlsgaard, & Roberts, 2000; Thal et al., 2004); children with hearing impairments (Tomblin, Spencer, Flock, Tyler, & Gantz, 1999); children with specific language impairment (SLI; Hadley, 1998a; Hewitt, Hammer, Yont, & Tomblin, 2004; Rice, Redmond, & Hoffman, 2006); and children who have been diagnosed with other clinical conditions such as autism, Down syndrome, and Fragile X syndrome (Price et al., 2008; Rollins & Snow, 1998; Scarborough, Rescorla, Tager-Flusberg, Fowler, & Sudhalter, 1991; Thordardottir, Chapman, & Wagner, 2002). For example, Hadley (1998a) used IPSyn to track the morpho-syntactic development of 20 mainstream English–speaking children with SLI when they were between the ages of 2 and 3 years. Although the children's IPSyn scores increased as they aged, their scores were significantly below those of the typically developing children at every testing session. For mainstream English, these findings show IPSyn to be sensitive to both developmental changes within children and developmental differences between children as a function of their clinical status.

Another study that is relevant to the current work was completed by Hewitt et al. (2004). Within this study, IPSyn scores of 6-year-olds with SLI were compared to those of typically

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

developing controls. Although the participants in this study exceeded the intended age limit of the IPSyn, Hewitt et al. argued that the tool contains some late-developing syntactic structures, and the scores of Scarborough's 4-year-olds were not at the tool's ceiling. Given this, Hewitt et al. reasoned that there was sufficient evidence to warrant an exploratory study of its use with older children. These authors also argued that this type of study was needed because the field lacks language sample measures of grammar for children older than 4 years. We agree with these authors, especially because grammar deficits have repeatedly been shown to be robust phenotypic markers of childhood language impairment (for review and a recent study, see Rice, Taylor, & Zubrick, 2008; for additional commentary about the role of grammar in the study of childhood language impairment, see Tager-Flusberg & Cooper, 1999).

To complete the IPSyn study, Hewitt et al. (2004) used language samples from fifty-four 6-year-olds (27 with SLI and 27 age-matched controls) who had participated in a large epidemiology study of SLI. Although there were language samples from 216 children in the original data set, the authors limited their analyses to those from White monolingual speakers of English. Samples from children of other ethnic and racial groups were excluded because of concerns about the children's potential uses of nonmainstream English. Also, Hewitt et al. included all of the items on the IPSyn except those found on the question/negation subscale. Items from this subscale were excluded because the authors found limited representation of questions within the language samples. Even with this subscale excluded, group differences were detected in the children's IPSyn sentence structure subscale scores and IPSyn total scores. As predicted, the scores of the children with SLI were lower than those of the controls, and effect sizes of these group differences were in the medium range (sentence structure subscale: $\eta^2 = .11$; total: $\eta^2 = .12$).

Finally, three IPSyn studies have been completed with children who speak nonmainstream dialects of English (Horton-Ikard et al., 2005; Oetting, 2005; Oetting et al., 1999). Oetting et al. (1999) and Oetting (2005) calculated children's IPSyn scores twice, once with utterances containing nonmainstream features included within the language samples and once with these same utterances removed. In Oetting et al., the participants were fifty-three 4- and 6-year-olds who spoke a rural variety of southern White English (SWE); in Oetting, the participants were forty 4- and 6-year-olds who spoke a rural variety of southern AAE. Across both studies, results showed that the children's IPSyn scores were unaffected by the presence of nonmainstream English features within the samples. In the third study, Horton-Ikard et al. (2005) examined children's IPSyn scores using data from twenty-two 2- and 3-year-olds who were learning a northern variety of AAE. The analyses involved correlations between the children's IPSyn scores and their AAE use, and AAE use was based on the rates (i.e., densities) at which the children produced nonmainstream English features within their utterances. Results showed the relationship between the children's IPSyn scores and their densities of nonmainstream English to be low in magnitude and statistically not significant.

Findings from these three nonmainstream dialect studies suggest that IPSyn is a dialect-neutral measure of children's use of grammar. Interestingly, though, in the two studies that included 4- to 6-year-olds, IPSyn was not always sensitive to changes in the children's grammars as a function of their age. Oetting et al. (1999) showed effects for age, with scores of the 6-year-olds higher than those of the 4-year-olds, but similar age effects were not found in Oetting (2005). Also, in both the Oetting et al. and Oetting studies, 6-year-olds with and without SLI were included as participants, and in both, the IPSyn scores did not vary as a function of the children's clinical status. These findings are inconsistent with those of Hewitt et al. (2004), who also studied 6-year-olds. These mixed findings warrant additional scrutiny, especially because the participants' dialects within these studies differed. The participants studied by Hewitt et al. spoke a mainstream dialect of English, those studied by

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Oetting et al. spoke a nonmainstream southern White dialect of English, and those studied by Oetting spoke AAE. If it turns out that the IPSyn scoring system can be used to detect differences between children as a function of their age and/or clinical status within some, but not all, dialects of English, then this would dampen our enthusiasm for the measure. In fact, this type of finding would lead us to abandon our view of IPSyn as dialect neutral. This would also mean that as a field, we would need to modify the IPSyn scoring system and/or develop an entirely new language sample measure to better meet the needs of all children.

The current study was designed to provide more information about these mixed findings and to further evaluate the clinical usefulness of IPSyn for nonmainstream English speakers by using AAE-speaking 4- and 6-year-olds as participants. To do this, we first completed an item analysis to determine if all of the items included within the IPSyn scoring system are produced by AAE-speaking children. If any of the items were not found within the children's samples and their omissions were related to the children's dialect, then regardless of the results from the children's total scores, we would want clinicians to remove these items from the IPSyn scoring system. Thus far, research on the IPSyn system has focused on children's total scores and/or their subscale scores. These types of studies do not provide clinicians with information about the appropriateness of the individual items on the tool.

Next, we completed a series of group comparisons to examine the children's IPSyn scores as a function of four variables. The first two variables, maternal education and nonmainstream dialect density, were related to the cultural and linguistic backgrounds of the AAE-speaking children. Evaluating the children's IPSyn scores as a function of their maternal education seemed important to do because previous studies have shown this variable to be correlated to some measures of children's development of language (e.g., mean length of utterance [MLU] and standardized vocabulary test scores; Dollaghan et al., 1999) and to children's use of nonmainstream English (Pruitt, 2006; see also Washington & Craig, 1998, for similar findings using family income level instead of maternal education as an index of poverty). Finding a difference in children's IPSyn scores as a function of their maternal education level would not necessarily lead to a rejection of this measure for clinical purposes, but it would indicate the need for alternative norms or a large normative sample to accommodate these differences. Finding a difference in the children's IPSyn scores as a function of their nonmainstream dialect density would lead us to question the dialect-neutral nature of the measure.

The third and fourth variables we examined as part of the group comparison analysis were the children's age (4 vs. 6) and clinical status (± SLI). Recall that the literature has been mixed as to whether IPSyn, when used with 4- and 6-year-olds, can be used to detect differences between children's grammar abilities as a function of these two variables.

This study addressed the following research questions.

- Are all of the IPSyn items appropriate for children who speak AAE?
- Do AAE-speaking children's IPSyn scores vary as a function of their maternal education level and/or nonmainstream dialect density?
- Is IPSyn sensitive to age-related changes in children's grammar for AAE speakers between the ages of 4 and 6 years?
- Do the IPSyn scores of AAE-speaking 6-year-olds with and without SLI differ from each other?

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

METHOD

Participants

Sixty-two African American children living in urban/suburban areas in southeastern Louisiana contributed data to the analyses as part of their participation in two recent dissertation studies (Garrity, 2007; Pruitt, 2006). In Garrity (2007), the focus was on the clinical condition of childhood SLI; in Pruitt (2006), the focus was on the language abilities of children reared in poverty as indexed by maternal education. Both of these studies were approved by our university's Institutional Review Board, with caregiver consent obtained before data collection.

The data included language samples from 32 typically developing 6-year-olds (6N); 15 of the children were classified as presenting low vocabulary ability and low socioeconomic status (LSES), and 17 were classified as presenting average vocabulary ability and middle socioeconomic status (MSES). Low vocabulary was determined by a standardized score ≤ 90 on the Peabody Picture Vocabulary Test—Third Edition (PPVT–III; Dunn & Dunn, 1997); LSES was determined by a maternal education level of <12 years. All but 1 of these 15 children also attended a school where >90% of the children received free lunch and the school's average standardized test scores were below the state average. Average vocabulary ability was determined by a standardized score > 90 on the PPVT–III, and MSES was determined by a maternal education level of at least high school. All but 4 of these 17 children attended schools where <10% of the children received free lunch and the school's average standardized test scores were above the state average. For both the LSES and MSES groups, the children did not have a history of speech and language services and their teachers were not concerned about their skills in these areas.

Ten additional samples were from 6-year-olds with SLI. SLI status was determined by receipt of services by a speech-language clinician and performance on a battery of tests: above −1 *SD* of the normative average on a nonverbal cognitive measure, which was the average score of the Figure Ground and Form Completion subtests of the Leiter International Performance Scale—Revised (LEITER–R; Roid & Miller, 1997), and below −1 *SD* of the normative mean on two language measures, the PPVT–III and three subtests that make up the syntax quotient of the Test of Language Development—Primary: Third Edition (TOLD–P:3; Hammill & Newcomer, 1997).

The 20 remaining language samples were from typically developing children who served as language matches within their respective studies; 15 were selected as vocabulary matches to the LSES group in Pruitt (2006), and 10 were selected as MLU matches to the SLI group in Garrity (2007). These children were recruited from preschools and child care facilities. The maternal education level of all but 2 of the children was ≥ 12 years, and the children's ages ranged from 47 to 66 months. We refer to these children by their average, median, and modal age in years, which was four (4N), but as can be seen, the ages of the children in this group were highly variable. The effect of this variability on the results is considered within our interpretation of the findings given that this type of within-group variability makes it difficult to detect statistically significant differences between groups.

Table 1 provides a breakdown of the children's maternal education levels and nonmainstream dialect ratings as a function of their group membership (6N, SLI, and 4N, with the 6N group further divided into LSES and MSES subgroups). As detailed in the original studies, the children's status as AAE speakers and their rates of nonmainstream English use was confirmed through blind listener judgments following the procedures of Oetting and McDonald (2002). Within this task, three raters independently listened to a 1-min excerpt of each child's language sample and then classified the child's dialect and rate

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

of nonmainstream English using a 7-point Likert scale. A score of 1 on the scale indicated that the rater perceived *no use of nonmainstream English,* and a score of 7 indicated *heavy use* (i.e., >40% of utterances). A child's nonmainstream dialect score was then calculated by averaging the listeners' ratings. The raters were certified speech-language pathologists who were also doctoral students in communication sciences and disorders and who had taken course work in nonmainstream dialects of English.

As an additional check of the listener judgments, we also calculated and report the children's nonmainstream dialect densities using a token-based method. As shown by Oetting and McDonald (2002), multiple methods exist for calculating a child's nonmainstream dialect density, and most of them are moderately to highly correlated to one another. The token-based measure we calculated was the percentage of the children's utterances that contained one or more nonmainstream English feature(s). As expected, the children's listener judgment ratings and their token-based dialect densities were significantly related to each other, $r = .68$, $p < .001$. Both measurement approaches also led to the same results within our analyses. Therefore, within the current work, we report our findings using the listener judgments given the ease at which a clinician could use a listener judgment method within clinical practice.

Table 1 also includes the children's test data and their MLU levels. To examine these data, one-way analyses of variance (ANOVAs) with group (LSES, MSES, SLI, 4N) as a between-subjects variable were completed. Group differences were detected for the children's dialect ratings, $F(3, 58) = 5.19$, $p = .003$, $\eta^2 = .21$; LEITER–R scores, $F(3, 58) = 3.63$, $p < .018$, $\eta^2 = .16$; PPVT–III scores, $F(3, 58) = 47.61$, $p < .001$, $\eta^2 = .71$; and TOLD–P:3 scores, $F(3, 58) = 28.13$, $p < .001$, $\eta^2 = .59$. Tukey follow-up testing of these differences indicated that the dialect ratings of the LSES group were higher than those of the MSES, SLI, and 4N groups, and the LEITER–R scores of the LSES group were lower than those of the 4N group. The PPVT–III and TOLD–P:3 scores of the LSES and SLI groups also were lower than those of the MSES and 4N groups, with the TOLD–P:3 scores of the SLI group also lower than those of the LSES group. Findings for the PPVT–III and TOLD–P:3 illustrate the apparent overlap that can exist between children reared in poverty who do not present the clinical condition of SLI and children with SLI when some standardized test scores are the focus of the comparison. Additional studies are needed to further explore these findings, but at a minimum, they underscore the need for the field to develop tools that can be used to better separate the language limitations of children reared in poverty from those with clinical conditions such as SLI (for probe data that address this issue, see Pruitt & Oetting, 2009).

Language Samples and IPSyn Scoring

Language samples were used to calculate the children's IPSyn scores. These samples were collected at each child's school, and they involved examiner–child interactions. Items used to facilitate conversation for the samples included a toy gas station, cars, people, picnic/park sets, Legos, a baby doll, baby care items, and three pictures portraying everyday activities (Arwood, 1985). The samples were transcribed using Systematic Analysis of Language Transcripts software (SALT; Miller & Iglesias, 2006). The total number of complete and intelligible utterances within the samples was 9,817, and the average number of complete and intelligible utterances per child was 158.34 ($SD = 52.57$). All but six of the samples were longer than 100 utterances. Those containing <100 utterances came from the typically developing 6-year-olds (4 LSES and 2 MSES), and they ranged in size from 55 to 97 utterances. Excluding these shorter samples from the analyses did not alter the results.

IPSyn scoring procedures followed the published guidelines of Scarborough (1990). Items in which two sufficiently different exemplars were identified within the samples were awarded a score of 2, and items in which one exemplar was identified were awarded a score of 1. If

NIH-PA Author Manuscript

an item was not found within a sample, it received a score of 0. The maximum score a child could earn was 112 (56 structures × 2) and maximum totals for the subscales were noun phrase = 22, verb phrase = 32, question/negation = 20, and sentence structure = 38. Both Computerized Profiling (Long, 1986) and SALT software facilitated manual searches for the 56 structures.

Reliability

Within the original studies, transcription reliability was examined at the utterance boundary and morpheme level using 10% of the samples. Across studies, interrater agreement for these measures was >90%. For the current study, the reliability of the children's IPSyn scores was assessed by having a second researcher independently calculate IPSyn scores for 10% of the samples. At the item level, interrater agreement was 93.5% (range = 91%–98%).

RESULTS

Item Analysis

The item analysis was done in two ways. First, we examined the IPSyn items against what is known about the morphosyntactic characteristics of AAE and other nonmainstream dialects of English using a list of 36 nonmainstream features that we have described elsewhere (35 are described in Oetting & McDonald, 2001, with the nonmainstream *go* copula added as a 36th feature in Oetting & Pruitt, 2005). For this analysis, IPSyn items were identified if they could be zero marked by an AAE speaker (e.g., *Yesterday he walk,* which includes a zero-marked form of past tense) or overtly marked with a nonmainstream form (e.g., *I ain't gonna do it,* which includes a nonmainstream auxiliary, and *He be funny,* which includes nonmain-stream habitual *be*). Our identification of these items was overly liberal and was based on the assumption that a clinician may not assign credit to overtly marked, nonmainstream expressions even though the IPSyn coding system never states that these options cannot receive credit. Following this analysis and using the children's full language samples, we calculated the percentage of AAE-speaking children who earned a 0 on each IPSyn item. For this analysis, we explored the full samples rather than restricting the children's samples to 100 utterances because our question focused on whether the items on IPSyn are (or are not) produced by AAE-speaking children. Given the consistency of the results across children, we report these analyses collapsed across groups.

As shown in Table 2, we deemed 17 (30%) of the 56 IPSyn items to have the potential to receive a 0 if a child produced a nonmain-stream English feature for the item 100% of the time. Of these 17 items, eight (47%) are part of the 11-item IPSyn verb phrase subscale. Following each item listed on this table, the percentage of children who actually earned a 0, 1, or 2 is also listed. As can be seen, the scoring of only two of these items, Q8 (which requires the production of an inverted modal, copula, or auxiliary within yes/no questions) and S17 (which requires an infinitive clause with a new subject), led to 50% or more of the children earning a score of 0.

Visual inspection of the data showed that some of the children who earned a 0 on Q8 produced utterances that could not be credited because of their use of nonmainstream English; either these children produced the question without inversion (e.g., *I'm coming tomorrow?*) or they zero marked the modal, copula, or auxiliary (e.g., *She coming tomorrow?*). Both of these types of questions are felicitous in AAE and other nonmainstream dialects of English. For S17, however, we did not find a single infinitive clause with a new subject that could not be credited because of a zero-marked infinitive *to* (e.g., *I want you help me*), which is another felicitous feature of AAE. Instead, children earned a 0 on this item because they did not produce utterances with infinitive clauses that

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

contained new subjects (e.g., children produced *I want to help* but not *I want you to help*). To our knowledge, avoidance of infinitive clauses with new subjects is not a feature of AAE or any other nonmainstream dialect of English that is spoken in the South. Given this, the children's low scores on this structure could not be linked to their use of non-mainstream English.

For all other items listed in Table 2, one or two sufficiently different exemplars of each structure were found in most of the children's samples. For readers who are familiar with AAE or other non-mainstream dialects of English, these findings should not be too surprising. This is because speakers of these dialects typically produce nonmainstream expressions in an optional or variable manner (e.g., 10%, 20%, 30%, etc.) rather than categorically (i.e., 100% of the time). The current set of results reflects this characteristic of nonmainstream English with AAE-speaking children who are 4 to 6 years of age.

Table 3 presents a list of the IPSyn items that received a score of 0 by 50% or more of the children studied. Our use of 50% as a criterion was relatively arbitrary; the data naturally divided at this cutoff, and we reasoned that if 50% or more of the children did not produce the item, then this would provide sufficient evidence to consider removing the item from the measure when working with AAE-speaking children. Interestingly, the resulting list contained only seven items, and none of these items came from the IPSyn verb subscale. Also, none of these items received a 0 by 100% of the children. This indicates that every item targeted on IPSyn was produced by at least 1 child. That only seven items appear on this list also means that the other 49 items (or 88% of the total items) were produced by >50% of the children. In addition, 53 (95%) of the items were produced by at least 40% of the children, and only one item (i.e., use of tag questions) was produced by <10% of the children. Finally, of the seven items that were produced by <50% of the children, the scoring of only Q8 could be linked to the children's use of nonmainstream English.

Group Comparisons

Table 4 lists the children's IPSyn totals and subscale scores by their group membership (6N, SLI, 4N, with the 6N group further divided into LSES and MSES subgroups). Given that some of the children's samples were longer than the recommended 100 utterances, the table contains two sets of IPSyn scores: The first set was generated with all of the complete and intelligible utterances within the samples, and the second set was generated with no more than 100 utterances within each sample. This was done because children's IPSyn scores have been found to be affected by sample size (Scarborough, 1990). Preliminary examination of our data also showed that for all groups, the children's IPSyn scores were higher when longer samples as opposed to 100-utterance samples were used, $F(1, 61) = 85.05$, $p < .001$, $\eta^2 = .58$, but the two sets of IPSyn scores resulted in the same statistical outcomes for all of our planned group comparisons. Given this, we report the statistical results using the children's IPSyn scores from the full samples. Full samples were selected to be consistent with the three previous IPSyn studies of children in the 4- to 6-year-old age range (i.e., Hewitt et al., 2004; Oetting, 2005; Oetting et al., 1999). Importantly, though, all across-study comparisons were completed with sample size differences taken into consideration. Also, all analyses were run with IPSyn totals and subscale scores; however, unless the results varied as a function of these different measures, only those findings generated from the children's IPSyn totals are reported.

IPSyn by maternal education and dialect density—To control for potential effects related to the children's age and clinical status, these analyses were completed with data from the 32 typically developing 6-year-olds. For these children, their maternal education levels were correlated to their nonmainstream dialect ratings, $r = -.36$, $p = .046$, but a

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

significant relationship was not found between their maternal education levels and their IPSyn scores, $r = -.03$, $p = .87$, or between their dialect ratings and their IPSyn scores, $r = .20$, $p = .27$. The children's IPSyn scores also did not vary as a function of their LSES and MSES groupings, $F(1, 30) = .02$, $p = .88$, $\eta^2 = .00$, or as a function of their dialect ratings when we used a median split to divide the children into low and high nonmainstream dialect speakers, $F(1, 30) = -1.59$, $p = .23$, $\eta^2 = .05$. These findings are consistent with those of the correlational analyses, and together they show the IPSyn scoring system to be unaffected by the children's maternal education levels and use of nonmainstream English.

IPSyn by age—To control for potential effects related to the children's clinical status, these analyses included data from the 52 children who were developing language typically. Correlations between the children's ages in months and their IPSyn totals and noun, verb, and sentence structure subscale scores were negligible (all $r < .20$, $p > .05$). We were also unable to detect a statistically significant group difference between the children's IPSyn totals and the noun, verb, and sentence structure subscales as a function of their ages (4N vs. 6N). For the children's question/negation subscale, the results were different. For this subscale, the children's ages were negatively correlated to their scores, $r = -.42$, $p = .002$, and the average scores of the 4N group were higher than those of the 6N group, $F(1, 50) = 10.49$, $p = .002$, $\eta^2 = .17$. These results are counterintuitive, but they are consistent with Hewitt et al.'s (2004) report that the 6-year-olds they studied did not produce questions within their samples.

To further explore this finding, we calculated the rates of questions that were produced by our 4- and 6-year-olds. For this analysis, all utterances with a question mark, including declaratives with rising intonation, were counted. Results showed the rates of questions produced by the 6-year-olds to be very low (5%), and these rates were lower than the 4-year-olds' rates of questions (13%), $F(1, 50) = -10.61$, $p = .002$, $\eta^2 = .18$.

IPSyn by clinical status—Given that the children's IPSyn scores did not vary as a function of their maternal education, nonmainstream dialect rate, or age, these analyses involved all of the children's scores. This led to a comparison between the scores of the 10 children with SLI and the scores of the 52 children without SLI. Results showed that the children's IPSyn scores were not found to vary by their clinical status, $F(1, 60) = .05$, $p = .82$, $\eta^2 = .001$. This null finding remained even when we restricted the comparison to the 6N and SLI groups, $F(1, 40) = .30$, $p = .59$, $\eta^2 = .007$, or to the MSES and SLI groups, $F(1, 25) = .47$, $p = .50$, $\eta^2 = .018$. Moreover, only the children's scores on the IPSyn sentence structure subscale were significantly correlated to any of their other language measures. This measure was the children's MLU, $r = .48$, $p < .001$.

DISCUSSION

The current study was conducted to learn more about the validity of using IPSyn for children who speak AAE. To examine this issue, we completed an item analysis and a comparison of the children's IPSyn scores as a function of their maternal education, nonmain-stream dialect density, age, and clinical status. Results showed that most of the items included within IPSyn are appropriate targets for children who speak AAE. In fact, we would argue that only one of the items (i.e., Q8, which involves the production of an inverted modal, copula, or auxiliary within yes/no questions) presents real potential for being negatively affected by a child's use of nonmain-stream features associated with AAE. We characterize this potential as real because 55% of the AAE-speaking children studied earned 0 points on this item, and this item could be linked to the children's use of nonmainstream English.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Although 50% or more of the children also earned 0 points for six other IPSyn items (i.e., N11, Q9, Q10, S9, S17, and S18), these items could not be directly linked to the children's use of non-mainstream English. Given this, a likely explanation for the children's low scores on these items is a limited opportunity to produce these structures within examiner–child conversation. Some evidence to support this interpretation can be found in studies by Eisenberg (1997, 2003), Fey and Leonard (1984), Craig and Washington (1994), and Jackson and Roberts (2001). All five of these studies examined children's use of language during conversational samples. In Eisenberg's studies, infinitives with new subjects were infrequently produced by children; in Craig and Washington's and Jackson and Roberts' studies, tag questions and gerunds were infrequently produced by children; and in Fey and Leonard's study, children's rates of questions within their samples decreased with age. Recall that we also found children's rates of questions to decrease with age in our study.

The results of the current study also showed that the children's IPSyn scores did not vary as a function of their maternal education levels or their nonmainstream dialect densities. This finding complements results from the item analysis by showing that the IPSyn scoring system is not systematically biased against some groups of AAE speakers as compared to others. Unfortunately, the children's IPSyn scores also did not vary as a function of their ages. Recall that in our previous studies, effects for age have been inconsistent. In Oetting et al. (1999), 6-year-olds earned higher IPSyn totals and sentence structure scores than 4-year-olds, but in Oetting (2005), an age effect was not detected. In the former, the participants were speakers of SWE; in the latter, they were speakers of AAE. Although the participants' dialects within these two studies were not the same, their average IPSyn scores were within 1 point of each other when the dialect groups were compared to each other (i.e., 4-year-olds: SWE = 84 vs. AAE = 86; 6-year-olds: SWE = 89 vs. AAE = 92). Given these IPSyn scores (and the repeated finding of no relationship between the children's nonmainstream dialect densities and their IPSyn scores), the unreliable effects for age that we have documented for IPSyn do not appear to be tied to the children's particular type of English dialect. Instead, we interpret these unreliable effects for age across studies as showing a limitation of IPSyn when it is used with children who are older than 48 months.

To further examine this issue, Table 5 presents a comparison of the children's IPSyn scores across studies. To facilitate the comparison, we organized scores by the number of utterances (75 vs. 100 vs. 100+) within the children's samples. When comparing these studies to each other, it is important to note that the ages of the 4-year-olds within these studies are not identical. Scarborough's 4-year-olds were 48 months of age, whereas those of the other 4-year-old groups were more variable (mean age in current study = 57 months; mean age in Oetting et al., 1999 = 48 months; mean age in Oetting, 2005 = 56 months). In spite of the age variation that is observed across the 4-year-old groups, the data in this table show that across dialects, developmental changes in IPSyn are evident in the 2- to 4-year age range, with plateau after the age of 48 months. This finding supports the use of IPSyn for AAE-speaking children, but it also suggests that for all children, this measure has an upper age limit. This finding is consistent with cautionary statements about IPSyn's upper age limit in Scarborough (1990). This upper age limit also makes sense to us from a child acquisition perspective because IPSyn is a measure of morphosyntactic emergence and not a measure of productivity or mastery.

Finally, within the current study, we were unable to detect a difference between the IPSyn scores of children with and without SLI. This finding (and our inability to detect other group differences within the data) cannot be readily explained by small and unequal sample sizes. This is because the correlational analyses also showed the children's IPSyn scores to be unrelated to most of the other measures we collected. Recall that none of the children's IPSyn scores (totals and subscale scores) was positively related to their ages, and only the

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

children's MLUs (and not their TOLD–P:3 scores) were correlated at a low level to their IPSyn sentence structure scores. In our opinion, far more correlations and stronger correlations should have been evident in the data for the null findings to be explained by limited statistical power of the analyses.

The lack of an observed difference between the children with and without SLI is consistent with Oetting et al. (1999) and Oetting (2005), but is inconsistent with the findings of Hewitt et al. (2004). Table 6 presents a comparison of the children's IPSyn scores from the current study and those from Hewitt et al. For this comparison, IPSyn scores from the children's full samples are used because Hewitt et al. also used some samples that included >100 utterances. As can be seen, the children's scores from the two studies are highly similar. In fact, all of the SLI versus typically developing group differences (between and within studies) appear well within the ranges of the standard deviations that accompany the means. Moreover, within the Hewitt et al. study, 30% of the children with SLI earned an IPSyn score that was above the typically developing children's average IPSyn score. Together, these findings confirm our earlier concerns about the clinical sensitivity (or lack thereof) of IPSyn at the 6-year-old age range.

Not surprisingly, Hewitt et al. (2004) also questioned the clinical sensitivity of IPSyn for 6-year-olds in spite of finding differences between children with and without SLI on the IPSyn sentence structure subscale and total. When discussing their results, Hewitt et al.'s concerns primarily focused on the limits of conversational samples for eliciting advanced grammar structures. We share their concern because some of IPSyn's advanced grammar structures were minimally represented within our samples. In addition to Hewitt et al.'s concern, we raise an additional concern that relates to IPSyn's focus on only two tokens of each structure. Although the measurement of two tokens for each structure appears appropriate for documenting the emergence of morphosyntax in children under the age of 48 months, for children over this age, measurement of more tokens is needed to determine whether a child is presenting dialect-appropriate levels of productivity and mastery.

For nonmainstream English-speaking children, perhaps the more important point to be drawn from Table 6 is that the children's scores across studies are similar even though the children spoke different dialects of English. This finding complements results from the item analysis and further attests to the appropriateness of the IPSyn scoring system for children who speak AAE. Given that other non-mainstream dialects of English contain many of the same nonmain-stream grammatical features of AAE, we also believe that the current set of findings will generalize to other samples of nonmainstream English-speaking children. Additional studies are needed to test this hypothesis directly.

Implications for Future Research

Although the current set of findings shows IPSyn to be a dialect-neutral assessment tool for measuring children's use of grammar, there are other types of studies that need to be done with the items and scoring procedures to further test its validity. The most obvious is the need for a large cross-sectional and multidialectal study of items on the measure using children who are between the ages of 2 and 4 years. This type of study would allow researchers to evaluate the diagnostic sensitivity and specificity of the items using children whose ages are ideally suited for this work. This type of work could focus on all of the items on IPSyn or on a theoretically motivated subset of items, as has been advocated by Hadley (1998a; for cross-linguistic and cross-etiological studies that support this type of approach, see Levy & Schaeffer, 2003; Oetting & Hadley, 2008; Rice & Warren, 2004).

Multidialectal studies that are longitudinal in nature are also needed. This type of work is necessary to document children's developmental growth trajectories of grammar across

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

different dialects of English. Longitudinal studies that can be used to guide this work include Hadley and Short (2005) and Hadley and Holt (2006). The focus of these two studies has been on young children's emergence, productivity, and eventual mastery of a subset of items that appear on the IPSyn verb phrase subscale. Findings from the current study illustrate the feasibility of extending this type of longitudinal research to children who speak nonmainstream English. This type of work would not only allow us to determine the types of grammar structures that show parallel growth or different growth across dialects, but within a dialect, it would also allow us to evaluate the impact of different rates (e.g., low, medium, high) of nonmainstream dialect use on children's acquisition of these structures.

Implications for Clinical Practice

The findings from this study provide support for the use of IPSyn as a language sample measure for AAE-speaking children because we did not find the items on the tool to be inappropriate targets for the dialect of AAE. We also did not find visually detectable differences between the IPSyn scores of the AAE-speaking children studied here and those previously reported for mainstream English-speaking children. These are important findings because IPSyn focuses on children's development of grammar, and valid assessment measures of grammar for nonmainstream English–speaking children are sorely missing within the field. In spite of these findings, we do not recommend IPSyn for children who exceed the age of 4 years (i.e., 48 months). This recommendation applies to children regardless of their English dialect. As we mentioned earlier, one reason IPSyn should not be used for children older than 48 months relates to its focus on grammar emergence without consideration of dialect-appropriate levels of productivity or mastery. The second reason relates to IPSyn's reliance on a conversational sample for eliciting complex grammar structures. As discussed by Hewitt et al. (2004) and as shown by our data, limited opportunity exists for children to use some of IPSyn's complex grammar structures during examiner–child conversation. Limitations of conversational samples for eliciting complex syntax are well documented in the school-age literature (Eisenberg et al., 2008; Nippold, Hesketh, Duthie, & Mansfield, 2005). The current findings show that these limitations are also relevant to children in the 4- to 6-year-old age range. Given this, we recommend that clinicians who work with children over the age of 48 months consider language sampling protocols that elicit text-level discourse from children (for a sampling protocol designed for this purpose, see Hadley, 1998b) and elicitation probes of complex grammar (e.g., Craig & Washington, 2002; Schuele & Tolbert, 2001). We recommend these types of tools to accompany dialect-neutral, norm-referenced tests that also target grammar (e.g., Diagnostic Evaluation of Language Variation—Norm Referenced Test; Seymour, Roeper, & de Villiers, 2005).

Acknowledgments

Appreciation is extended to the children who provided the data for the study and to the caregivers and teachers who supported the research. Appreciation is also extended to Lesli Cleveland, Heidi Huckabee, and Beth Wooden, who helped with the elicitation and transcription of the language samples, and to Michael Hegarty, Ann Mulkern, and Janet McDonald, who helped with the IPSyn coding criteria and/or the analyses. For 15 of the participants studied here, preliminary IPSyn scores were calculated as part of an MA thesis by Emily Jones.

References

Arwood, EL. Apricot I language kit. Portland, OR: Apricot; 1985.

Craig H, Washington J. The complex syntax skills of poor urban, African-American preschoolers at school entry. Language, Speech, and Hearing Services in Schools. 1994; 25:181–190.

Craig H, Washington J. Oral language expectations for African American preschoolers and kindergarteners. American Journal of Speech-Language Pathology. 2002; 11:59–70.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Dollaghan CA, Campbell TF, Paradise JL, Feldman HM, Janosky JE, Pitcairn DN, Kurs-Lasky M. Maternal education and measures of early speech and language. Journal of Speech, Language, and Hearing Research. 1999; 42:489–501.

Dunn, LM.; Dunn, LM. Peabody Picture Vocabulary Test. 3. Circle Pines, MN: AGS; 1997.

Eisenberg S. Investigating children's language: A comparison of conversational sampling and elicited production. Journal of Psycholinguistic Research. 1997; 26:519–538.

Eisenberg S. Production of infinitival object complements in the conversational speech of 5-year-old children with language impairment. First Language. 2003; 23(3):327–341.

Eisenberg SL, Ukrainetz TA, Hsu JR, Kaderavek JN, Justice LM, Gillam RB. Noun phrase elaboration in children's spoken stories. Language, Speech, and Hearing Services in Schools. 2008; 39:145–157.

Fey ME, Leonard LB. Partner age as a variable in the conversational performance of specifically language-impaired and normal-language children. Journal of Speech, Language, and Hearing Research. 1984; 27:413–423.

Garrity, AW. Unpublished doctoral dissertation. Louisiana State University; Baton Rouge: 2007. A study of auxiliary BE in African American English: A comparison of children with and without specific language impairment.

Hadley PA. Early verb-related vulnerability among children with specific language impairment. Journal of Speech, Language, and Hearing Research. 1998a; 41:1384–1397.

Hadley PA. Language sampling protocols for eliciting text-level discourse. Language, Speech, and Hearing Services in Schools. 1998b; 29:132–147.

Hadley PA, Holt JK. Individual differences in the onset of tense marking: A growth curve analysis. Journal of Speech, Language, and Hearing Research. 2006; 49:984–1000.

Hadley PA, Short H. The onset of tense marking in children at risk for specific language impairment. Journal of Speech, Language, and Hearing Research. 2005; 48:1344–1362.

Hammill, D.; Newcomer, P. Test of Language Development—Primary. Vol. 3. Austin, TX: Pro-Ed; 1997.

Hewitt LE, Hammer CS, Yont KM, Tomblin JB. Language sampling for kindergarten children with and without SLI: Mean length of utterance, IPSyn, and NDW. Journal of Communication Disorders. 2004; 38:197–213. [PubMed: 15748724]

Horton-Ikard R, Ellis Weismer S, Edwards C. Examining the use of mainstream language production measures in the language samples of African-American toddlers. Journal of Multilingual Communication Disorders. 2005; 3:169–182.

Jackson SC, Roberts JE. Complex syntax production of African American preschoolers. Journal of Speech, Language, and Hearing Research. 2001; 44:1083–1096.

Lee, L. Developmental sentence analysis. Evanston, IL: Northwestern University Press; 1974.

Levy, Y.; Schaeffer, J. Language competence across populations: Toward a definition of specific language impairment. Mahwah, NJ: Erlbaum; 2003.

Long, SH. Computerized Profiling (Version 1.0) [Computer software]. Arcata, CA: Author; 1986.

Miller, J.; Iglesias, A. Systematic Analysis of Language Transcripts (SALT), English and Spanish (Version 9) [Computer software]. University of Wisconsin—Madison; Waisman Center: Language Analysis Laboratory; 2006.

Nelson, NW.; Hyter, YD. How to use Black English sentence scoring (BESS) as a tool of non-biased assessment. Short course presented at the American Speech-Language-Hearing Association; Seattle, WA. 1990 Nov.

Nippold MA, Hesketh LJ, Duthie JK, Mansfield TC. Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. Journal of Speech, Language, and Hearing Research. 2005; 48:1048–1064.

Oetting, JB. Assessing language in children who speak a non-mainstream dialect. In: Ball, M., editor. Clinical sociolinguistics. Oxford, England: Blackwell; 2005. p. 180-192.

Oetting JB, Cantrell JP, Horohov JE. A study of specific language impairment (SLI) in the context of non-standard dialect. Clinical Linguistics and Phonetics. 1999; 13:25–44.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Oetting, JB.; Hadley, PA. The morphosyntax of children with language impairments. In: Schwartz, RG., editor. The handbook of child language disorders. New York, NY: Psychology Press; 2008. p. 341-364.

Oetting JB, McDonald J. Nonmainstream dialect use and specific language impairment. Journal of Speech, Language, and Hearing Research. 2001; 44:207–223.

Oetting JB, McDonald J. Methods for characterizing participants' nonmainstream dialect use in child language research. Journal of Speech, Language, and Hearing Research. 2002; 45:508–518.

Oetting JB, Pruitt SL. Southern African American English use across groups. Journal of Multilingual Communication Disorders. 2005; 3:136–144.

Price JR, Roberts JE, Hennon EA, Berni MC, Anderson KL, Sideris J. Syntactical complexity during conversation of boys with Fragile X syndrome and Down syndrome. Journal of Speech, Language, and Hearing Research. 2008; 51:3–15.

Pruitt, SL. Unpublished doctoral dissertation. Louisiana State University; Baton Rouge: 2006. Grammatical morphology of children reared in poverty: Implications for specific language impairment.

Pruitt SL, Oetting JB. Past tense marking by African American English-speaking children reared in poverty. Journal of Speech, Language, and Hearing Research. 2009; 52:1–14.

Rescorla L, Dahlsgaard K, Roberts J. Late-talking toddlers: MLU and IPSyn outcomes at 3;0 and 4;0. Journal of Child Language. 2000; 27:643–664. [PubMed: 11089342]

Rice ML, Redmond SM, Hoffman L. Mean length of utterance in children with specific language impairment and in younger control children shows concurrent validity and stable and parallel growth trajectories. Journal of Speech, Language, and Hearing Research. 2006; 49:793–808.

Rice ML, Taylor CL, Zubrick SR. Language outcomes for 7-year-old children with and without a history of late emergence at 24 months. Journal of Speech, Language, and Hearing Research. 2008; 51:394–407.

Rice, ML.; Warren, S. Developmental language disorders: From phenotypes to etiologies. Mahwah, NJ: Erlbaum; 2004.

Rispoli M, Hadley P. The leading edge: The significance of sentence disruptions in the development of grammar. Journal of Speech, Language, and Hearing Research. 2001; 44:1131–1143.

Roid, G.; Miller, L. Leiter International Performance Scale—Revised. Chicago, IL: Stoelting; 1997.

Rollins P, Snow C. Shared attention and grammatical development in typical children and children with autism. Journal of Child Language. 1998; 25:653–673. [PubMed: 10095329]

Scarborough HS. Index of Productive Syntax. Applied Psycholinguistics. 1990; 11:1–22.

Scarborough HS, Rescorla L, Tager-Flusberg H, Fowler AE, Sudhalter U. The relation of utterance length to grammatical complexity in normal and language-disordered groups. Applied Psycholinguistics. 1991; 12:23–45.

Schuele CM, Tolbert L. Omission of obligatory relative markers in children with specific language impairment. Clinical Linguistics and Phonetics. 2001; 15:257–274.

Seymour, HN.; Roeper, T.; de Villiers, J. Diagnostic Evaluation of Language Variation—Norm Referenced Test. San Antonio, TX: The Psychological Corporation; 2005.

Stockman I. Promises and pitfalls of language sample analysis as an assessment tool for linguistic minority children. Language, Speech, and Hearing Services in Schools. 1996; 27:355–366.

Tager-Flusberg H, Cooper J. Present and future possibilities for defining a phenotype for specific language impairment. Journal of Speech, Language, and Hearing Research. 1999; 42:1275–1278.

Thal DJ, Reilly J, Seibert L, Jeffries R, Fenson J. Language development in children at risk for language impairment: Cross-population comparisons. Brain and Language. 2004; 88:167–179. [PubMed: 14965539]

Thordardottir ET, Chapman RS, Wagner L. Complex syntax production by adolescents with Down syndrome. Applied Psycholinguistics. 2002; 23:163–183.

Tomblin JB, Spencer L, Flock S, Tyler R, Gantz B. A comparison of language achievement in children with cochlear implants and children using hearing aids. Journal of Speech, Language, and Hearing Research. 1999; 42:497–511.

Washington JA, Craig HK. Dialectal forms during discourse of poor, urban African American preschoolers. Journal of Speech and Hearing Research. 1994; 37:816–823. [PubMed: 7967568]

Washington JA, Craig HK. Socioeconomic status and gender influences on children's dialectal variations. Journal of Speech, Language, and Hearing Research. 1998; 41:618–626.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 1

Characteristics of the study participants—32 typically developing 6-year-olds (6N), 15 classified as presenting low vocabulary ability and low socioeconomic status (LSES) and 17 classified as presenting average vocabulary ability and middle socioeconomic status (MSES); ten 6-year-olds with speech-language impairment (SLI); and 20 typically developing children who served as language matches within their respective studies (4N).

		Age (months)			Maternal education[a]			Dialect rating[b]			DDM[c]			LEITER–R[d]			PPVT–III[e]			TOLD–P:3[f]			MLU[g]		
	N	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range
6N	32	73.09	3.68	67–83	12.78	2.98	6–16	4.85	1.31	1.7–7.0	27	15	7–45	10.08	1.73	6.0–15.0	91.72	12.97	68–117	90.81	14.86	59–124	6.49	1.21	4.22–8.99
LSES	15	73.47	4.02	67–83	10.00	1.41	6–11	5.58	1.03	3.0–7.0	32	10	12–45	9.47	1.55	6.0–11.5	80.27	6.60	68–90	81.13	15.27	59–124	6.49	1.38	4.22–8.99
MSES	17	72.76	3.44	67–81	15.24	1.35	12–16	4.22	1.21	1.7–6.0	24	10	7–41	10.63	1.74	8.5–15.0	101.82	7.50	89–117	99.35	7.65	89–115	6.48	1.09	4.48–8.13
SLI	10	77.50	4.77	66–85	11.56	1.42	10–15	5.20	.74	4.0–6.7	36	08	21–49	10.30	1.51	9.0–13.5	78.00	4.90	72–87	66.10	8.49	57–79	5.10	.89	3.78–6.81
4N	20	57.55	5.21	47–66	14.85	1.79	10–16	4.47	1.18	2.0–6.3	26	10	7–49	11.48	2.13	9.0–16.0	99.30	7.29	88–114	99.65	10.73	83–121	5.64	1.02	3.84–7.06

[a]Years of education completed by the child's mother; all but 1 participant provided this information.

[b]Mean listener judgment rating of African American English (AAE) use; maximum score = 7, with higher scores indicating higher ratings of nonmainstream AAE.

[c]Dialect density measure, calculated as the percentage of utterances the child produced with one or more nonmainstream English structure(s).

[d]Average score on the Figure Ground and Form Completion subtests of the Leiter International Performance Scale—Revised (LEITER–R; Roid & Miller, 1997); *M* = 10, *SD* = 3.

[e]Standard score on the Peabody Picture Vocabulary Test—Third Edition (PPVT–III; Dunn & Dunn, 1997); *M* = 100; *SD* = 15.

[f]Syntax quotient on the Test of Language Development—Primary: Third Edition (TOLD–P:3; Hammill & Newcomer, 1997); *M* = 100; *SD* =15.

[g]Mean length of utterances reported in morphemes.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Index of Productive Syntax (IPSyn; Scarborough, 1990) items with the potential to be influenced by nonmainstream English.

IPSyn item	Percentage of children earning 0	Percentage of children earning 1	Percentage of children earning 2
N4 nominal preceded by article or modifier	0	0	100
N5 article used before noun	0	0	100
N6 two-word noun phrase after verb/preposition	0	0	100
N7 plural suffix	0	0	100
V4 copula linking two nominals	0	0	100
V6 auxiliary *be*, *do*, *have* in verb phrase	0	3	97
V7 progressive suffix	0	0	100
V10 third person singular present tense suffix	24	34	42
V12 regular past tense	3	11	86
V13 past tense auxiliary	2	16	82
V15 copula, modal, or auxiliary for emphasis/ellipsis	39	21	40
V16 past tense copula	5	48	47
Q6 wh question with inverted modal, copula, auxiliary	32	37	31
Q7 negation of copula, modal, or auxiliary	2	6	92
Q8 Y/N question with inverted modal, copula, auxiliary	55	22	23
S8 infinitive without catenative, marked with *to*	0	5	95
S17 infinitive clause with new subject	58	27	15

Note. The items are listed and numbered by their respective IPSyn subscale: N = noun phrase subscale, V = verb phrase subscale, Q = question/negation subscale, and S = sentence structure subscale.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Table 3

IPSyn items with 50% or more of the children earning a score of 0.[a]

IPSyn item	Percentage of children earning 0	Percentage of children earning 1	Percentage of children earning 2
N11 any other bound morpheme on noun or adjective	55	22	23
Q8 Y/N question with inverted modal, copula, auxiliary	55	22	23
Q9 *why, when, which, whose*	63	26	11
Q10 tag question	92	6	2
S9 *let/make/help/watch* introducer	55	12	32
S17 infinitive clause with new subject	58	27	15
S18 gerund	77	12	11

Note. The items are listed and numbered by their respective IPSyn subscale: N = noun phrase subscale, V = verb phrase subscale, Q = question/negation subscale, and S = sentence structure subscale.

NIH-PA Author Manuscript

Table 4

IPSyn scores by sample size and group.

	Total[a]		Noun[b]		Verb[c]		Question/negation[d]		Sentence structure[e]	
	M	SD	M	SD	M	SD	M	SD	M	SD
Full samples										
6N	87.69	10.96	20.25	1.32	27.31	2.91	10.94	4.76	29.19	4.28
LSES	88.00	12.92	20.20	1.42	27.40	3.29	11.27	4.86	29.13	5.11
MSES	87.41	9.29	20.29	1.26	27.24	2.64	10.65	4.80	29.24	3.56
SLI	89.70	6.45	20.70	1.06	27.50	3.24	12.80	2.94	28.70	2.58
4N	91.10	5.24	20.05	1.0	27.90	2.55	14.65	2.37	28.50	2.78
100-utterances samples										
6N	81.91	9.04	19.78	1.04	25.13	3.15	9.50	4.16	27.50	4.18
LSES	80.40	10.35	19.53	.99	24.93	3.26	9.13	3.72	26.80	4.57
MSES	83.24	7.77	20.00	1.06	25.29	3.14	9.82	4.60	28.12	3.84
SLI	81.90	8.57	19.80	1.55	23.80	4.29	11.60	3.10	26.70	3.34
4N	83.75	5.39	19.55	1.0	24.70	3.0	13.35	2.64	26.15	2.98

[a]Maximum score = 112;

[b]Maximum score = 22;

[c]Maximum score = 32;

[d]Maximum score = 20;

[e]Maximum score = 38.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 5

Across-study comparison of IPSyn totals by sample size and child age.

Sample size and study	2½ years		3½ years		4 years		6 years	
	M	SD	M	SD	M	SD	M	SD
75 utterances								
African American English, Horton-Ikard et al. (2005)	52.27	5.16	70.19	8.52	—	—	—	—
Mainstream English, Scarborough (1990)[a]	52.73	10.13	72.20	7.23	80.53	5.48	—	—
100 utterances								
African American English, Current study	—	—	—	—	83.75	5.39	81.91	9.04
Mainstream English, Scarborough (1990)[a]	58.80	10.67	79.53	6.37	85.80	4.21	—	—
100+ utterances								
African American English, Current study	—	—	—	—	91.10	5.24	87.69	10.96
Southern White English, Oetting et al. (1999)	—	—	—	—	84.00	5.67	89.63	5.23
African American English, Oetting (2005)	—	—	—	—	85.92	9.49	91.75	11.69

[a]This study reflects a longitudinal design; others listed are cross-sectional.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 6

Across-study comparison of 6-year-olds' IPSyn scores by dialect.

	Total minus question/negation[a]		Noun[b]		Verb[c]		Sentence structure[d]	
Dialect	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Nonmainstream English (Current study)[e]								
SLI (*n* = 10)	76.90	5.33	20.70	1.06	27.50	3.24	28.70	2.58
Typically developing (*n* = 32)	76.75	7.31	20.25	1.32	27.31	2.91	29.19	4.28
Mainstream English (Hewitt et al., 2004)								
SLI (*n* = 27)	74.48	7.64	20.11	1.18	27.33	3.86	26.77	3.93
Typically developing (*n* = 27)	79.14	5.08	20.51	.89	29.00	2.56	29.40	3.55

[a]Maximum score = 92;

[b]Maximum score = 22;

[c]Maximum score = 32;

[d]Maximum score = 38.

[e]Scores calculated using each child's full samples to best align with the samples of Hewitt et al. (2004).