Summer 8-8-2017

Language Assessment in African American English-Speaking Children: A Review of the Literature Since 1983 and Grammaticality Judgments of Low-Income, African American English-Speaking Children: The Role of Language Ability and Dialect Density

Ryan Lee

ACCEPTANCE

This dissertation, LANGUAGE ASSESSMENT IN AFRICAN AMERICAN ENGLISH-SPEAKING CHILDREN: A REVIEW OF THE LITERATURE SINCE 1983 AND GRAMMATICALITY JUDGMENTS OF LOW-INCOME, AFRICAN AMERICAN-SPEAKING CHILDREN: THE ROLE OF LANGUAGE ABILITY AND DIALECT DENSITY, by RYAN MARIE LEE, was prepared under the direction of the candidate's Dissertation Advisory Committee. It is accepted by the committee members in partial fulfillment of the requirements for the degree, Doctor of Philosophy, in the College of Education and Human Development, Georgia State University.

The Dissertation Advisory Committee and the student's Department Chairperson, as representatives of the faculty, certify that this dissertation has met all standards of excellence and scholarship as determined by the faculty.

_____
Julie A. Washington, Ph.D.
Committee Chair

_____
Robert Hendrick, Ph.D.
Committee Member

_____
Laura Fredrick, Ph.D
Chairperson, Department of Educational Psychology, Special Education, & Communication Disorders

_____
Debra McKeown, Ph.D.
Committee Member

_____
Cynthia Puranik, Ph.D
Committee Member

_____
Date

_____
Paul A. Alberto, Ph.D.
Dean
College of Education & Human Development

AUTHOR'S STATEMENT

By presenting this dissertation as a partial fulfillment of the requirements for the advanced degree from Georgia State University, I agree that the library of Georgia State University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote, to copy from, or to publish this dissertation may be granted by the professor under whose direction it was written, by the College of Education and Human Development's Director of Graduate Studies, or by me. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without my written permission.

_____

Ryan Marie Lee

NOTICE TO BORROWERS

All dissertations deposited in the Georgia State University library must be used in accordance with the stipulations prescribed by the author in the preceding statement. The author of this dissertation is:

Ryan Marie Lee
Department of Educational Psychology, Special Education, and Communication Disorders
College of Education and Human Development
Georgia State University

The director of this dissertation is:

Julie A. Washington
Department of Educational Psychology, Special Education, and Communication Disorders
College of Education and Human Development
Georgia State University
Atlanta, GA 30303

CURRICULUM VITAE

Ryan Marie Lee

ADDRESS: 3305 Wesley Stonecrest Circle
Lithonia, GA 30038

EDUCATION:

Ph.D.	2017	Georgia State University Department of Educational Psychology, Special Education, & Communication Disorders
M.S.	2011	West Virginia University Speech-Language Pathology
B.A.	2009	Michigan State University Communication Sciences & Disorders

PROFESSIONAL EXPERIENCE:

2014-2017	Data Collection/Reduction Supervisor Learning Disabilities Research Innovation Hub Georgia State University
2012-2014	Speech-Language Pathologist Early Intervention-Early Steps Baton Rouge, LA
2011-2014	Graduate Research Assistant Language Development and Disorders Lab
2012	Louisiana State University Independent Study Speech Acoustic Lab Louisiana State University

PRESENTATIONS AND PUBLICATIONS:

Lee-James, R., & Washington, J. A. (in preparation). Language assessment in African American English-Speaking children: A review of the literature since 1983.

Lee-James, R., & Washington, J. A. (in preparation). Grammaticality judgments of low-income, African American English-speaking children: The role of language ability and dialect density.

Washington, J. A., Branum-Martin, L. A. Sun, C., & **Lee-James, R.** (under review). "Is dialect density a marker of risk?: Evidence from language and reading.".

Washington, J. A., Branum-Martin, **Lee-James, R.,** Sun, C. (in preparation). Reading and language performance of low-income, African American boys in grades 1 – 5.

Lee-James, R., Washington, J. A., & Seidenberg, M. (submitted). Morphological judgments of low-income, school-age African American English-speaking children.

Lee, R., & Oetting, J. (2014). Zero marking of past tense in child African American English. *SIG I: Perspectives on Language Learning and Education, 21,* 173-181.

Oetting, J., **Lee, R.,** & Porter, K. (2013). Evaluating the grammars of children who speak nonmainstream dialects of English. *Topics in Language Disorders, 33,* 1-12.

Johnson, L., & **Lee James, R.,** Gatlin, B., & Patton-Terry, N. (2016 November). *Executive functioning skills and its relation to oral and written narrative skills in African American children.* Poster presented at the American Speech-Language-Hearing Association Convention, Philadelphia, PA.

Lee James, R., & Washington, J. A. (June 2015). *Performance of low-income, African American English-speaking children on the Morphological Comprehension subtest of the TOLD-I:4.* Poster presented at the Symposium on Research in Child Language Disorders, Madison, WI.

Lee, R., Oetting, J., & Hegarty, M. (June 2013). *Past tense and past participle structures in child AAE and SWE.* Poster presented at the Symposium on Research in Child Language Disorders, Madison, WI.

Gregory, K., **Lee, R.,** & Oetting, J. (April 2013). *Working with students who speak AAE & SWE: What SLPSs should know*. Seminar presented at the National Black Association for Speech, Language and Hearing, Washington, D.C.

Oetting, J., Riviere, A., **Lee, R.,** Richardson, J., Hegarty, M., & McDonald, J. (June 2012). *Structure specific differences across tasks in nonmainstream dialects.* Poster presented at the Symposium on Research in Child Language Disorders, Madison, WI.

PROFESSIONAL SOCIETIES AND ORGANIZATIONS

2017	American Speech-Language-Hearing Association
2014	Southern Region Education Board
2014	National Black Association for Speech-Language and Hearing
2013	Sigma Xi, The Scientific Research Society

LANGUAGE ASSESSMENT IN AFRICAN AMERICAN ENGLISH-SPEAKING CHILDREN: A REVIEW OF THE LITERATURE SINCE 1983

by

RYAN LEE JAMES

Under the Direction of Julie A. Washington

ABSTRACT

The overarching purpose of both dissertation studies is to contribute to the extant literature base on language assessment in the context of poverty and African American English (AAE) dialect. Language assessment with culturally and linguistically diverse populations, in particular children who speak AAE, has been a longstanding challenge for professionals in the field of speech-language pathology despite the preponderance of scholarly attention this topic has received. The purpose of the first study is to conduct a systematic literature review to synthesize the existing literature on AAE from the past approximately three and a half decades, to identify the aspects of language and assessment approaches that have been most informative for identifying language impairment in this population. The purpose of the second study is to examine the grammaticality judgments of school-age, AAE-speaking children as a function of their nonmain-

stream dialect density and language ability. Data for this study came from 273 African American children from low-income backgrounds who were participants in a larger project focused on language and literacy outcomes for children reared in urban areas. The relationship between language ability and dialect density was explored using correlational analysis and the contribution of language ability and dialect density on grammaticality judgments was analyzed using multiple regression. Finally, a multivariate analysis of variance was conducted to investigate the impact of dialect density and language ability on various items that differed in grammatical constructions. Results from both studies are discussed relative to the existing oral language profiles of AAE speakers and the impact of linguistic variation on assessment. Together, these papers contribute to the extant literature by supporting the development of a more comprehensive profile of AAE and increasing the field's understanding of language assessment and language impairment in child AAE speakers.

INDEX WORDS: African American English, Dialect, Language assessment, Language impairment, Language development.

GRAMMATICALITY JUDGMENTS OF LOW-INCOME, AFRICAN AMERICAN ENGLISH-SPEAKING CHILDREN: THE ROLE OF LANGUAGE ABILITY AND DIALECT DENSITY

by

RYAN MARIE LEE

A Dissertation

Presented in Partial Fulfillment of Requirements for the

Degree of

Doctor of Philosophy

in

Students with Exceptionalities

in

Department of Educational Psychology, Special Education, and Communication Disorders

in

the College of Education and Human Development

Georgia State University

Atlanta, GA

2017

DEDICATION

To the children and families I have worked with throughout the years who have been the cornerstone of my learning.

ACKNOWLEDGMENTS

All that I have been able to accomplish would not have been possible without the grace and strength of my Lord and Savior, Jesus Christ, and the sacrifices of my ancestors. To my mentor and role model, Dr. Julie A. Washington, I am forever indebted to you for adopting me, advising me, and helping me reach my academic and professional goals. To my esteemed committee, your guidance and support has been instrumental in my growth as a scientist and I grateful that you have invested in me. To the people and organizations who made my doctorate financially possible—Garry Bingham, Nicole Patton-Terry, Julie Washington, the Georgia Learning Disabilities Research Hub, the Language and Literacy Initiative, and the Southern Regional Education Board—thank you for investing my dreams. Last but certainly not least, I could not have achieved this without the unconditional love and support of my husband, Ray James, my beautiful daughter, Kindle Lee James, my parents, Richard and Stacy Lee, and my loving aunt, Lynn Lee.

TABLE OF CONTENTS

LIST OF TABLES

ABBRIVIATIONS

AAE	African American English
AA	African American
LI	Language Impairment
SLI	Specific Language Impairment
SES	Socioeconomic Status
DELV-ST	Diagnostic Evaluation of Language Variation-Screening Test
DELV-NR	Diagnostic Evaluation of Language Variation-Norm Referenced
TOLD-I:4	Test of Language Development-Intermediate:4th Edition
MC	Morphological Comprehension Subtest of the TOLD-I:4
CTOPP-2	Comprehensive Test of Phonological Processing-2nd Edition

1 LANGUAGE ASSEMENT IN AFRICAN AMERICAN ENGLISH-SPEAKING CHILDREN: A REVIEW OF THE LITERATURE SINCE 1983

Difference versus Disorder: Background and Statement of the Problem

Language assessment in culturally and linguistically diverse populations has historically posed a challenge for professionals in the field of speech-language pathology (Stockman, 1996; Stockman, 2010). A population of particular interest has been children who speak African American English (AAE), a major dialect of American English that is differentiated from Mainstream American English (MAE) in all language domains—but most significantly in the use of morphosyntax (Seymour, Bland-Stewart, & Green, 1998). Much of this research has focused on African American (AA) children growing up in poverty, as income status has been observed to impact the degree to which children use dialect (Craig & Washington, 2004; Washington & Craig, 1998), with low-income children using higher levels of dialect than their middle income peers. Children reared in poverty may also present with language differences that are not associated with dialect. For example, many of these children have limited vocabulary and shorter mean length of utterances (Dollaghan et al., 1999), making it difficult to dissociate the effects of income status and dialect use. Language differences attributed to AAE use and language difference attributed to socio-cultural factors, namely socioeconomic status (SES) have been the source of diagnostic difficulty and a chief focus of research for decades. As whole, the findings from this work reveal how language differences have the potential to contribute to underestimation of African American (AA) children's language skills, especially on standardized assessments (Cole & Taylor, 1990; Hammer, Pennock-Roman, & Rzasa, 2002; Kresheck & Nicolosi, 1973; Rodekohr & Haynes, 2001; Stockman, 1996; Thomas-Tate, Washington, Craig, & Pack-

ard, 2006; Washington & Craig, 1992; Washington & Craig, 1999). This dilemma has been called the *difference vs. disorder* controversy. The *difference versus disorder* debate can be linked to the fields of sociolinguistics, education, and communication disorders. The earliest impressions of "Negro Dialect" were in the field of sociolinguistics and perpetuated a disordered perspective (Fortier, 1884). As a consequence, speaking a nonmainstream dialect is often associated with negative stereotypes of the speakers (Wolfram, Temple Adger, Christian, 1999). This was true of AAE until researchers actively sought to dispel beliefs about its legitimacy. Data-based studies on AAE served to legitimize the dialect by documenting the productions of adult and adolescent speakers, establishing AAE as a rule-governed, oral language system rather than a deficient form of English (Fasold, 1969; Labov, Cohen, Robins, & Lewis, 1968; Rickford, 1999; Rickford & Theberge-Rafal, 1996; Wolfram, 1969; Wolfram & Fasold, 1972; Wolfram & Thomas, 2002). These studies provided support for three landmark education decisions, commonly referred to as Lau v. Nichols, the Ann Arbor Decision, and the Oakland Ebonics Debate (Lau v. Nichols, 1974; Martin Luther King Junior Elementary School Children et al. v. Ann Arbor School District, 1979; Wolfram, 1998). These resolutions advanced the field of education by acknowledging inequity in the quality of education provided for AA students, advocating for "Black-English" or "Ebonics" to be recognized as AA children's first language, and advising educators to use the dialect to support and facilitate learning in public school settings.

In the field of speech-language pathology *difference versus disorder* has been used to represent the complex set of challenges faced by clinicians and researchers when assessing children who speak nonmainstream dialects of English and often by extension, children from low SES backgrounds. The Position Statement on *Social Dialects* developed by the American Speech and Hearing Association (ASHA, 1983) was designed to address this issue in the disci-

pline by maintaining that "...no dialectal variety of English is a disorder or a pathological form of speech or language", and recognizing that "..each social dialect is adequate as a functional and effective variety of English". Unfortunately, neither the Position Statement, nor the court cases provided professionals with specific resources or strategies for serving children from diverse backgrounds (Seymour, 2004). However, ASHA's stance on linguistic diversity did provide a platform for researchers to study and address challenges in clinical practice and this research has led to a greater understanding of the impact of AAE in language development, language impairment, and literacy. Studies focused on AAE have begun to unpack the *difference versus disorder* dilemma in a number of ways, including documenting features of child AAE (Craig, Thompson, Washington, & Potter, 2003; Washington & Craig, 1994), providing rate-based data for the frequency of productions of hallmark features (Craig & Washington, 2004; Newkirk-Turner, Oetting, & Stockman, 2014; Oetting & McDonald, 2001; Ross, Oetting, & Stapleton, 2004; Roy, Oetting, & Moland, 2013), exploring dialect variation in the context of socioeconomic status (Craig, Washington, Thompson-Porter, 1998; Craig & Washington, 2002; Washington & Craig, 1998; Pruitt & Oetting, 2009; Horton-Ikard & Miller, 2004), examining dialect classification and measurement indices (Oetting & McDonald, 2002; Horton-Ikard & Apel, 2014; Terry, Connor, Thomas-Tate, & Love, 2010), documenting variability in dialect use across tasks (Oetting et al., 2012; Thompson, Craig, & Washington, 2004; Washington, Craig, & Kushmaul, 1998) and studying the impact of dialect use on literacy and literacy-related skills (Apel & Thomas-Tate, 2009; Charity, Scarborough, & Griffin, 2004; Connor & Craig, 2006; Craig, Thompson, Washington, & Potter, 2004; Craig & Washington, 2004; Gatlin & Wanzek, 2015; Terry, 2006; Terry, 2012; Terry & Scarborough, 2011).

Recently, Oetting et al. (2016a) suggested replacing the phrase *dialect (or difference) versus disorder* with *disorder within dialect* in order to move away from phrasing which perpetuates that the two are mutually exclusive. The goal of the change in phrasing is to begin to endorse a perspective that childhood language impairment exists within and across dialects of English. In this view, foregrounding *disorder* allows the training and expertise of the speech-language pathologist (SLP) in diagnosis and treatment of language disorders to be the primary focus. For example, Oetting et al. (2016a) suggest that commonalities in language impairment across dialects should be the focus of a *disorder within dialect* framework including (a) positive family history as a risk factor, (b) importance of universal screenings in early identification of language impairment, and, (c) importance of using diagnostic indices to select fair and appropriate assessments. While re-wording the phrase to more accurately reflect the current state of understanding is a first step to improving service delivery for underrepresented children, it is argued here, that this change will have limited impact until clinicians have access to appropriate assessments, in particular assessments that are normed on AA children. Assessment challenges have been documented extensively in the extant literature and in response to these challenges scholars have identified potential solutions. This literature review presents empirical research evidence on language and assessment approaches that have been most informative for distinguishing *difference* from *disorder* in AAE speakers.

Purpose Statement

Despite the recent increase in the number of studies focused on *difference versus disorder*, inconsistencies in the methodologies, the clinical profile of participants, and concerns about the validity of assessment measures have led to an uncertain picture concerning which language skills and assessment approaches are most informative for identification of language impairment

in AAE speakers. The purpose of the current systematic review is to synthesize the literature on the topic of *difference versus disorder* in order to present a coherent clinical description of language impairment in AAE. Further, this literature reviews aims to synthesize the literature on AAE and language disorders from the past approximately three and a half decades, in an effort to identify the various language and assessment approaches that have been most informative for identifying language impairment in this population.

Studies published from 1983 to the present were selected for inclusion as this year represents a shift in the approach to the assessment and treatment of social dialects marked by the publication of ASHA's Position Statement on *Social Dialects*. The findings of this systematic review should serve to inform clinical practice directly and outline future research directions leading to improved services for AAE speakers.

Search Procedures, Inclusionary Criteria, and General Study Characteristics

The current literature review presents research evidence from 19 child language studies that met the following criteria: (a) focused on AAE-speaking, school age children (enrolled in pre-k through 5th grade or average age between three- to 12-years-old) with typical language ability and at least one clinical group with either language/speech delay or language/speech impairment determined a priori; (b) collected and included empirical oral language data; and (c) quantitative analysis methods were used to compare performance between clinical groups. Focusing on children ages three to 12-years old was important to capture the aspects of language crucial for diagnosis of language impairment as early childhood is a critical time for language acquisition (Brown, 1973) which supports the development of language and literacy skills throughout middle childhood (Catts, Fey, Zhang, & Tomblin, 1999; Storch & Whitehurst, 2002). Importantly, underrepresentation of African and language-minority children in early in-

tervention and special education, particularly in the categories of speech-language impairments and learning disabilities (Morgan et al., 2015; Morgan et al., 2016), underscores the need to improve language assessment procedures for AA children, who are often difficult to diagnose. Further, it was also important that these investigations were empirical and included clinical comparisons with speech/language impairments in order to increase their replicability and clinical utility. Of the 1,507 located in the search, 77 remained after the title search and were narrowed to 25 after reviewing the abstracts. Eighty-three studies were eliminated that did not meet the specific criteria for this review.

The 19 studies that met the criteria for this review were located by conducting a query into several databases and content-specific journals including *Academic Search Complete, ERIC, PHYSCArticles, PSYCHInfo, CINAHL* and *Journal of Speech, Language, and Hearing Research; Language, Speech, and Hearing Services in Schools; American Journal of Speech-Language Pathology; Journal of Communication Disorders; International Journal of Language and Communication Disorders;* and *Clinical Linguistic and Phonetics* using the following search terms: *Black English, Ebonics, African American Vernacular English, nonmainstream dialect, language impairment, language delay, and language disorders.*

Overall, each of the 19 studies included employed cross-sectional research designs, were conducted in various regions of the United States including the Midwest, Southeast, Northeast and the West, and included child participants from varying SES backgrounds. In general, these studies (Burns, de Villiers, Pearson, & Champion, 2012; Cleveland & Oetting, 2013; Craig & Washington, 2000; de Villiers, Roeper, Bland-Stewart, & Pearson, 2008; Garrity & Oetting, 2010; Johnson & de Villiers; 2009; Oetting et al., 2010; Oetting et al., 2016b; Oetting & Cleveland, 2006; Oetting & McDonald, 2001; Oetting & Newkirk, 2008; Rodekohr & Haynes, 2001;

Seymour et al., 1998; Washington & Craig, 2004) used similar criteria to define clinical groups, as follows: generally, children were selected for inclusion in language impairment groups (henceforth LI or SLI for specific language impairment) on the basis of the following criteria: (a) performance within normal limits on a measure of nonverbal intelligence (i.e., Columbia Mental Maturity Scale, Kaufman Assessment Battery for Children, Leiter International Performance Scale, Primary Test of Nonverbal Intelligence), (b) a previous diagnosis of language impairment by an SLP and receipt of speech-language therapy at the time of study, and (c) performance below -1 standard deviation (*SD*) of the mean on a single measure or battery of norm-referenced and/ or criterion-referenced language measures (e.g., Test of Language Development, Clinical Evaluation of Language Fundamentals, Oral and Written Language Scales, Peabody Picture Vocabulary Test, Diagnostic Evaluation of Language Variation-Screening Test, Diagnostic Evaluation of Language Variation-Norm Referenced, and mean length of utterances). Three studies were an exception (Laing, 2003; Rodekohr & Haynes, 2001; Wilcox & Anderson, 1998) as they failed to document whether their participants had normal intelligence and three studies (Burns et al., 2012; de Villiers et al., 2008; Johnson & de Villiers, 2009) reported all participants had normal intelligence however the assessment(s) used or performance data was not reported. Four studies included participants with intelligence scores at a cutoff of -2 *SD* of the mean rather than -1 *SD* (Stockman, 2008; Stockman, Guillory, Seibert, & Boult, 2013; Stockman, Newkirk-Turner, Swartzlander, & Morris, 2016).

There were some studies that arbitrarily assigned children to clinical groups. For example, Laing (2003), classified Head Start children as having speech delays if they were receiving speech-language services and typical if they were "judged to be developing normally" (p. 275) and did not provide information regarding who provided participant reports (e.g., parents, teach-

ers, SLPs) and what informal or formal measures were collected to assess speech and language skills. Wilcox and Anderson (1998) classified children as being typically developing (TD) or having a phonological delay on the basis of teacher report and an audiotaped speech sample that was analyzed by both authors. In order for children to be put in either the typical or impaired group both authors had to reach agreement. Wilcox and Anderson (1998) do not provide or discuss in detail a specific coding system that was used in their phonological analysis. All studies documented that participants had hearing within normal limits.

Some researchers included three clinical groups (Craig & Washington, 2000; Garrity & Oetting, 2010; Oetting et al., 2010; Oetting & McDonald, 2001; Oetting & Newkirk, 2008): one group with LI and two typical comparison groups—a group with typical language ability matched on age, and a group with typical language matched on language ability (e.g., using MLU, standardized measures of vocabulary and syntax). Although researchers have noted potential limitations to using a language-match group (Plant, Swisher, Kiernan, & Restrepo, 1993) the comparison between children with LI and younger children who are typical allows for a developmental comparison, or comparison of language levels, in that children with LI are often differentiated from language-match controls emphasizing the magnitude of the delay in linguistic ability.

The process of confirming and quantifying dialect status varied across studies. Some researchers used blind listener judgments (Cleveland & Oetting, 2013; Garitty & Oetting, 2010; Oetting et al., 2010; Oetting et al., 2016b; Oetting & Cleveland, 2006; Oetting & McDonald, 2001; Oetting & Newkirk, 2008)—a process that requires listeners to make judgments based on a speech sample and rate nonmainstream use on a Likert scale from 1 (no nonmainstream use) to 7 (very high nonmainstream use). Other researchers used a type-token based method, which in-

volves collecting and analyzing a language sample for target AAE structures (Craig & Washington, 2000; Laing, 2003; Oetting & McDonald, 2001; Rodekohr & Haynes, 2001; Seymour et al., 1998; Stockman, 2008a; Stockman, 2008b; Stockman et al., 2013; Stockman et al., 2016; Washington & Craig, 2004; Wilcox & Anderson, 1998). See Oetting and McDonald (2002) for a detailed discussion of blind listener judgments and Washington and Craig (1994, 1998) and Oetting and McDonald (2002) for discussions of type-token methods. Fewer researchers (Burns et al., 2012; Oetting et al., 2016b) used the Diagnostic Evaluation of Language Variation-Screening Test to classify dialect as either *Mainstream American English, Some Variation from MAE,* or *Strong Variation from MAE*. Johnson & de Villiers (2009) and de Villiers et al. (2008) classified children into dialect groups based on family and community background and the presence or absence of AAE features in children's responses.

Aspects of Language and Assessment Approaches Identified as Clinically Informative

Based on the outcomes of the selected studies, the aspects of language that have been most informative for language assessment within this population across domains can be sorted into three major categories: a) language domains that are absolutely critical to include in any language assessment with this population (i.e., expressive morphosyntax, phonology, and vocabulary), b) processing-dependent language skills designed to avoid the bias identified for knowledge dependent assessments (e.g., nonword repetition), and, c) general developmental language skills that when included are informative (i.e., language comprehension, and narratives, and sentence recall).

Critical Language Domains

The language domains in this section have received the most attention in research on *difference versus disorder*. The focus on these skills has confirmed their importance for inclusion in assessments. Two of these domains (i.e., morphosyntax and phonology) are influenced heavily by dialect and considered to be contrastive with MAE while vocabulary is not considered to vary between dialects but is more likely influenced by SES (Dollaghan et al., 1999; Hart & Risley, 1995; Washington & Craig, 1992). Further, vocabulary is unquestionably one of the earliest indicators of language and other developmental delays (Morgan et al., 2016).

Expressive morphosyntax. Expressive morphosyntactic aspects of language were identified most often as important to include in assessment. This is not a surprise given that AAE impacts morphosyntax significantly. Nine out of 19 studies examined children's expressive morphological and syntactic ability as clinical indicators of language impairment (Cleveland & Oetting, 2013; Craig & Washington, 2000; Garrity & Oetting, 2010; Oetting et al., 2010; Oetting & McDonald, 2001; Oetting & Newkirk, 2008; Seymour et al., 1998; Stockman et al., 2013; Stockman et al., 2016) and seven of these studies reported that various morphological and syntactic aspects of language successfully differentiated AAE speakers with and without (S)LI. As the exceptions, Cleveland and Oetting (2013) focused on verbal -s (also referred to as third person singular) and reported that six-year-old AAE speakers produced verbal *-s* at similar rates (TD = 21.42%, SLI = 14.07%) in spontaneous language samples, regardless of clinical status. Oetting et al. (2010) examined the clinical utility of extending the Index of Productive Syntax (IPSyn; Scarborough, 1990), an assessment tool developed for two- to four-year-olds, to assess the language skills of six-year-old AA participants and found that this tool was not sensitive to clinical status for children in this age group. Children with SLI (M = 89.70), age-matched controls (M = 87.69), and language-match controls (M = 91.10) performed comparably on the

overall measure. The IPSyn total score is comprised of four subscales that target noun phrases, verb phrases, question/negation, and sentence structure; scores on all four subscales were also comparable across clinical groups.

The remaining seven studies revealed significant performance differences between the (S)LI groups and groups of TD children. Overall these studies show divergence in the linguistic skills between groups on several aspects of language. Specifically, AAE-speaking children with (S)LI were found to have shorter mean length of communication units (MLCUs; Craig & Washington, 2000), produced limited amounts and types of complex syntax and attempted other non-contrastive features (i.e., present progressive, prepositions, and pronouns) less often (Craig & Washington, 2000; Oetting & Newkirk, 2008; Seymour et al., 1998), and zero marked certain grammatical morphemes (regular and irregular past, irregular third-person, and auxiliary BE, and non-inversion of Wh-question) at higher rates than their typical peers (Garitty & Oetting, 2010; Oetting & McDonald, 2001; Seymour et al., 1998). Head Start children with language delays studied by Stockman et al. (2013) and Stockman et al. (2016) failed to meet the criteria for the Minimal Competence Core-Morphosyntax which meant that they failed to produce one or more morphosyntactic features, in spontaneous speech, comprising a corpus of age- and dialect-appropriate features. Stockman and colleagues (2013, 2016) confirmed the validity of the MCC-MS by comparing performance of the pass/fail groups on measures of convergent validity (i.e., assessment tools expected to measure the same or similar constructs) such as the Preschool Language Scale-Third Edition (PLS-3; Zimmerman, Steiner, & Pond, 1992) and IPSYN. In both instances, children in the pass group earned significantly higher scores yielding robust effect sizes for group comparisons on the PLS-3 total score (d = 1.03) and the IPSYN (d = 2.96).

In the interest of practical and clinical relevance, certain details, commonalities, and findings that arose in these nine studies warrant further discussion. The first is that all nine studies used language sample analysis to assess expressive morphosyntax and while Garrity and Oetting (2010) developed a probe to elicit auxiliary BE, they used it in conjunction with language sample analysis. It is common in the field of speech-language pathology to employ language sample analysis as an assessment tool especially in research because this method boasts high ecological validity (Stockman, 1996). However, a documented drawback to language sample analysis is limited ability on behalf of the examiner to elicit complex syntax, grammatical constructions, or other targeted language behaviors (e.g., pragmatic functions such as conversational repairs) making probes and other cloze tasks a desirable alternative when the goal is to measure and compare language behaviors across ability groups. Expressive language probes and other cloze tasks are advantageous because of the more controlled structure. In these tasks, prompts and number of trials are consistent across participants, and language data are not dependent upon children's propensity to produce a given structure. Differences in elicitation methodology have the potential to lead to conflicting findings as can be seen in Garrity and Oetting (2010). When the dependent variable was collapsed BE productions (i.e., *am, is,* and *are*) from language sample data, no statistical differentiation between groups was detected, however on the elicitation probe a large effect for clinical status was established (partial η^2 = .19). The authors explain that there were limited BE tokens available for group comparison in the language sample condition which precluded the use of parametric statistics, resulting in the use of nonparametric methods which have less statistical power. Consequently, selecting one method versus another has the potential to impact language outcomes.

Another important finding is that certain morphological and syntactic elements of language differentiated groups and it is important to note that these elements were both contrastive (e.g., past tense and auxiliary BE) and noncontrastive (e.g., complex syntax). Contrastive features are ones that are used differently from MAE in particular omissions of grammatical morphemes (e.g., -ed) that are characteristic of the dialect but are obligatory in MAE. Noncontrastive features are shared between the two dialects. Although some scholars have discouraged the inclusion of contrastive features in language assessment with AA children because they overlap with clinical indicators of impairment in MAE (Seymour et al., 1998; Pearson, Jackson, & Wu, 2014), as presented here, certain contrastive aspects of grammar have the potential to be clinically informative. From these results it can be inferred that the term *contrastive* is fluid rather than absolute in that less prevalent AAE features such as zero marking of irregular and regular past tense are diagnostically salient whereas more prevalent features such as zero marking of third person singular are less likely to show production differences across clinical groups.

Finally, in four studies it was observed that dialect density was unrelated to children's performance (Garrity & Oetting, 2010; Oetting et al., 2010; Oetting & Newkirk, 2008; Stockman et al., 2013). No effect for dialect density confirms that certain aspects of language including complex syntax and MLCUs are indeed dialect-neutral. In other words, these aspects of language appear to be unimpacted by rate of nonmainstream dialect use making them ideal candidates for assessment. It is important to note that these studies mainly focused on noncontrastive aspects of language therefore reducing the potential influence of dialect density. The anomaly was the study focused on contrastive feature BE in which the correlation between dialect density and BE useage for the language-matched controls and SLI groups were nonsignificant whereas a strong, negative correlation was observed for typically-developing, age-match controls between

BE useage and dialect density (Garrity & Oetting, 2010). Overall this indicates that dialect density may be less informative for examining performance differences between or across groups with varying language ability, and more informative for examining trends in performance within children with typical language skills.

Phonology. The contrastive nature of AAE phonology has been a consistent area of inquiry for scholars interested in unpacking *difference versus disorder*. For the current literature review, three studies met the inclusionary criteria and focused on discriminating between TD groups and children diagnosed with phonological disorder using both standardized experimental assessments. Wilcox and Anderson (1998) employed an experimental measure, the Wilcox African-American English Screening Test of Articulation, to assess the speech sound repertoires of AA children through 50 words targeting seven word-initial singletons and 18 consonant clusters. Results showed that TD children were differentiated from children with impairment on the measure with a robust effect for clinical status (d = 2.40).

Laing (2003) reported a statistically significant decrease in final consonant deletion for both clinical groups using an alternative response mode. The data came from two standardized articulation tests (Goldman Fristoe Test of Articulation and the Analysis of Phonological Processes-Revised) and the stimulus were manipulated to include the word *eyes* at the end of each word (e.g., basket-eyes); thereby creating a linguistic environment sensitive to AAE phonotactic rules. Even with a significant reduction in final consonant deletion on both measures (d ranged from .62 to .78) for the group with speech delays, applying an alternative response method was not always enough to change their classification to TD. This was most likely due to productions of other noncontrastive phonological processes.

Finally, Stockman (2008) found that clinical status and percent consonants correct significantly predicted pass/fail outcomes on the Minimal Competence Phonetic Core (MCC-PH). To pass the MCC-PH, children needed to demonstrate phonological competence as evidenced by production of at least two different word-initial consonant clusters and a set of 13 pre-determined word-initial singletons; fail status was assigned to participants who did not produce one or more competency. There was a robust effect for pass/fail status (*d* exceeded 2.5) which indicated that percent of competencies met differed significantly between groups. Data for this study came from oral language samples in which the samples were analyzed for target consonants.

Taken together these studies showcase the importance of including noncontrastive elements of phonology (e.g., word-initial singletons and consonant clusters, fronting of velars, and stopping) when assessing AAE speakers. From a wider perspective, it appears that at least some popular standardized assessments, such as the Goldman Fristoe Test of Articulation may be appropriate for AA children with a critical caveat: the impact of contrastive phonological features (e.g., final consonant deletion or consonant cluster reduction) is controlled for in overall performance either through using an alternative response mode, or eliminating these items from analysis altogether. However, using measures that are intended for AAE speakers are likely to provide the most valid data because they were developed specifically for this population and by design should be inherently unbiased and dialect-sensitive. Although in the case of the current review these measures were experimental, Stockman (2008) showed that pass/fail status on the MCC-PH was significantly predicted by performance on a published criterion-referenced measure (Percent Consonants Correct-Revised; Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997) and the relationship between these two suggest that the MCC-PH may be a valid method for phonological assessment within this population.

Vocabulary. The study of vocabulary is in the semantic domain**.** Deficits in vocabulary are prevalent among children with LI (Morgan et al., 2016) however the diagnostic accuracy of standardized receptive and expressive vocabulary measures is questionable largely because of potential cultural bias (Champion, Hyter, McCabe, Bland-Stewart, 2003; Kresheck & Nicolosi, 1973; Restrepo et al., 2006; Stockman, 2000; Washington & Craig, 1992; Washington & Craig, 1999). In the current review, five studies used number of different words (NDW; Craig & Washington, 2000; Stockman, 2008; Stockman et al., 2013; Stockman et al., 2016; Washington & Craig, 2004), derived from oral language samples, as an index of expressive vocabulary. These investigations observed moderate to large effects for clinical status on this measure (Stockman, 2008 reported *d* = .98; Stockman et al., 2013 reported *d* = 1.37; Washington & Craig, 2004 reported η^2 = .12) and reported that NDW accounted for a significant portion (Stockman et al., 2016 reported r^2 = .18) of the variance in scores between TD children and those expected to have protracted language development. In a similar vein, an investigation examining children's ability to learn novel words, Johnson and de Villiers (2009) revealed a large effect for clinical status (partial η^2 = .16) equating to children with SLI having more difficulty assigning meaning to novel verbs across various argument structures including (a) intransitive (e.g., The boy is *temming.*), (b) transitive (e.g., The clown is *saiping* the woman.), (c) dative (e.g., The boy is *meeping* the flowers to the girl.) , and (d) complement (e.g., The woman is *ganning* the waiter to send coffee.). The measure used to assess word learning was developed and piloted for inclusion on the Dialect Sensitive Language Test (published as the DELV-NR). The task employed three pictures displaying actions taking place. The examiner read the stimulus sentence for example, "The boy is *temming*", and then asked the question "Which one is the *temmer*?". In order to answer successfully children need to be able to employ their knowledge of argument struc-

ture. Overall these studies show that vocabulary measures, in particular those that assess expressive vocabulary are informative for assessment with this population.

Processing-Dependent Measures

Unlike traditional knowledge-dependent measures that rely heavily on children's language knowledge (e.g., vocabulary, morphology, syntax), processing-dependent language measures reduce the likelihood of test bias occurring as a result of linguistic or cultural interference because these tasks rely on children's ability to process and manipulate linguistic information rather than accumulation of linguistic knowledge. A widely used processing-dependent measure, nonword repetition (NWR) tasks involves children repeating novel nonsense words that are presumably equally unfamiliar to all participants regardless of cultural background.

Investigating the clinical utility of NWR was the focus of three studies each of which followed the methods of Campbell, Dollaghan, Needleman, and Janosky (1997) and reported that NWR differentiated children with and without (S)LI who spoke AAE with scores favoring the TD group. For example, Washington and Craig (2004) reported that NWR and two other measures significantly contributed to the performance differences between pass/fail groups, accounting for 51% of the total variance in scores and the effect size for the group comparison on NWR was moderate (η^2= .145). In a NWR investigation conducted by Oetting and Cleveland (2006), results indicated that the accuracy for assigning clinical status was low when considered in isolation but improved when NWR was used in conjunction with other measures. In spite of the large effect for clinical status (partial η^2= .30) and a high level of overall classification accuracy (81%), specificity for the measure was extremely low (56%) yet when used in conjunction with a language comprehension measure, sensitivity and specificity increased to 81% and 94%, respectively.

In addition to NWR, Rodekohr and Haynes (2001) also used the Competing Language Processing Task, another processing-dependent measure that requires children to recall the last word of a sentence and this language measure in addition to NWR, was observed to differentiate children as well. An important observation of these studies was that participants in both clinical groups showed similarities in their repetition patterns. Specifically, scores decreased with increases in syllable length (one, two, three, and four syllables) suggesting that children with (S)LI as well as TD controls are impacted by increasing processing demands imposed by having to retain and repeat lengthening novel words though children with (S)LI lagged behind their TD peers in each syllable condition (Oetting & Cleveland, 2006; Rodekohr & Haynes, 2001).

General Language Skills

The language skills in this section have been included less often in assessment studies focused on AAE-speaking children, though when included, language comprehension, narrative ability, and sentence recall appear to be important indicators of LI. In early childhood comprehension and narrative skills, in particular, are indicative of language level and school readiness (Fiorentino & Howe, 2004; Pankratz, Plante, Vance, & Insalaco, 2007; Weismer, Branch, & Miller, 1994). Generally speaking, the literature reviewed reveals that these general aspects of language are important to evaluate in assessment as they broadly index more sophisticated language competencies that children must possess in order to be proficient language users.

Language comprehension. Two studies investigated children's comprehension of Wh-questions and observed that the LI group had difficulty answering Wh-questions. This difficulty was reflected by significant performance differences by clinical group and were qualitatively different in their response errors more often producing no-response and unrelated answers (Craig & Washington, 2000) in addition to answering *how* questions as *why* questions, answering false

clause questions incorrectly (e.g, *What did the mother say she bought?* where what the mother *bought* and *said she bought* are incongruent), and answering medial complementizer clause instead of the intended question (de Villiers et al., 2008). Craig and Washington (2000) used two picture stimulus to elicit Wh-questions ranging in difficulty from simple questions probing the participant to name objects (what questions) to more complex and cognitively taxing questions probing manner (how questions) and time (when questions). In the same way, de Villiers et al. (2008) used picture probe stimulus (items 13-22) from the Dialect Sensitive Language Test (published as the DELV-NR) to elicit responses to Wh-questions. Each question was accompanied by three stimulus pictures; children were told a short story about the pictures and then the examiner asked a one Wh-question in which children needed to use information provided in the stimulus pictures and in the story to answer.

Craig and Washington (2000) and Washington and Craig (2004) also reported data from another language comprehension measure that assessed children's understanding of active and passive sentences. Craig and Washington (2000) found that children with LI performed significantly lower than TD controls while Washington and Craig (2004) found no significant performance differences by group. These two studies differed in their targeted age groups in that one study included four- to 11-year-olds (Craig & Washington, 2000) and the other included four- to six-year-olds (Washington, & Craig, 2004). This data was elicited using a two-option, forced-choice picture selection format in which the examiner read active and passive sentences and children were asked to point to the picture that best represented the stimulus sentence.

Narrative ability. Assessment of four elements of narrative language including reference contrasting (i.e., distinguishing characters using noun and prepositional phrases), temporal expressions (i.e., using adverbial conjunctions such *next, then,* or *later*), mental state descriptions

(i.e., demonstrating theory of mind), and understanding behavior based on false beliefs (i.e., being able to inhibit one's own beliefs and identify the character's beliefs) substantiated group distinctions in narrative ability (Burns et al., 2012) as indicated by a, medium effect (η^2 = .13) on the composite index that included all four elements. Considered individually, results suggest that cohesion elements (i.e., reference contrasting and temporal expressions) are informative for children ages four to six though statistical performance differences were observed the effect was small between clinical groups (η^2 = .06); and this result was hampered by a significant interaction between age and clinical status indicating that differences between groups dissipated beyond age 6. In contrast, evaluation elements (i.e., mental state descriptions and understanding of false beliefs) of narrative ability were robust across ages four to nine for clinical status (η^2 = 1.0). As a whole, the findings of Burns et al. (2012) show that all four indices of narrative language are clinically informative and worth including in assessment with AA children although age of participants should be an important consideration. For this study, data were collected using two probes each of which consisted of six picture sequences. Children were asked to tell a story using the pictures as a guide. After the child completed their story the examiner asked questions to assess children's ability to infer mental states and understand false beliefs. These materials developed as a part of the Dialect Sensitive Language Test (published as the DELV-NR).

Sentence recall. In a study of nonmainstream dialect use, Oetting et al. (2016b) used a sentence imitation probe that consisted of pre-recorded sentences designed to assess recall of a one functional category: tense and agreement (e.g., Minnie *is* cleaning the dirty dishes in the sink.); two functional categories: tense and agreement and negation (e.g., Minnie *is not* cleaning the dirty dishes in the sink.); and three functional categories: tense and agreement, negation, and complement structures (e.g., Mickey wonders *if* Minnie *is not* cleaning the dishes.). Results

showed significant differences in group performance and the effect for clinical status was large (partial η^2 = .55) with TD children demonstrating greater accuracy in their ability to recall sentences. The clinical groups also differed qualitatively in their recalls with the SLI group producing more tense and negation errors than the TD controls; groups demonstrated comparable rates of errors on complements.

This study also examined the diagnostic accuracy of the sentence recall task by using three different scoring methods. The first method, was a 2-point scoring system that had been used by other researchers in the extant literature. This method assigns 2 points for exact recalls, 1 point for one to three errors, and no points are given if the child makes more than four errors. For this method, the maximum possible score was 72 and using a clinical cutoff of 40, this method yielded adequate levels of sensitivity (.89) and specificity (.86). The second method used 12 exact recalls as the cut point; children who produced less than 12 were classified as SLI. This method yielded almost the exact same sensitivity (.89) and specificity (.87). The final method involved using a cut point of .24 representing the proportion of ungrammatical recalls and yielded slightly lower levels of sensitivity (.86) and specificity (.80). The results of this study suggest sentence recall is clinically informative for distinguishing between TD AAE speakers and those with SLI.

Conclusion and Future Research Directions: What the Field Needs Now

The current literature review revealed several aspects of language that are clinically informative in assessment with children who speak AAE. Importantly, the aspects of language identified here are also those that have been identified as informative for language assessment with the general population of children who may or may not be speakers of a cultural dialect. The findings from this review highlight various aspects of language representative of a variety of

skills and language domains, while also revealing that the study of language assessment in AAE is lacking in depth and systematicity. In particular, these studies are lacking in quantity and some aspects of language have only been examined in certain age groups with the exception of expressive morphosyntax. Though approximately 30% (6/19) of the studies evaluated identified vocabulary or word learning ability as a clinical marker of language impairment, these studies were narrowly focused on younger children in preschool and kindergarten (Stockman, 2008; Stockman et al., 2013; Stockman et al., 2016; Washington & Craig, 2004); though two studies did include a children from ages four to 11 (Craig & Washington, 2000; Johnson & de Villiers, 2009). Similarly, the three studies on phonology only included children three- to six-years old. There were limited studies on language comprehension (3), processing-dependent measures (3), narrative ability (1), and sentence recall (1) and so although it is suggested that these language skills are clinically informative, perhaps it is most accurate to conclude that these skills and the measures used to access them appear to be informative for AA children in that particular age range. The field needs converging evidence to determine whether these aspects of language are indeed clinically useful and for whom (importantly considering age and SES). Expanding this area of inquiry to include wider age ranges should lead to the development of culturally and linguistically appropriate assessments normed on this population.

These studies also revealed a glaring absence of receptive language measures, with the exception of the three studies focused on responses to Wh-questions and passive/active sentences (Craig & Washington, 2000; Johnson & de Villiers, 2009; Washington & Craig, 2004). The remaining 13 studies were focused exclusively on children's expressive language abilities. Consequently, the current profile of AAE consists primarily of expressive language, though absence of receptive language data in this population was identified more than a decade ago as a critical

omission in the literature (Johnson, 2005). In an effort to create a more comprehensive profile of AAE, future studies should aim to examine comprehension of language structures that are both contrastive and noncontrastive with MAE. Prior research indicates that linguistic variability within language productions may impact the age of mastery for expressive tasks, with linguistic structures involving variable inclusion or exclusion emerging later for AAE-speaking children than for their MAE-speaking peers for whom these structures are obligatory (Thomas-Tate, Washington & Edwards, 2004). It is important to determine whether this variability exists within the receptive domains well.

To date, language development research in the context of AAE has been primarily cross-sectional. In fact, all 19 of the studies included in the current literature review employed a cross-sectional research design. This is unfortunate because longitudinal evidence would provide crucial information about how language develops over time in AAE speaking children, and how the relationship between dialect use and language ability changes as children grow older. This kind of longitudinal data would be invaluable for early identification, and perhaps prevention, of language and literacy disorders, and is critical for development of assessments and interventions needed for this population. Recently researchers have begun to examine AAE use longitudinally and the impact on literacy but this evidence is limited to children from kindergarten through second grade (Craig, Kolenic, & Hensel, 2014). Additional longitudinal evidence will be critical for improving the field's understanding of AAE as whole.

Finally, in addition to the focus on research, what the field desperately needs is to ensure that graduate programs across the country are training competent clinicians prepared to meet the needs of diverse populations specifically regarding the use of dynamic assessment. Of the 19 studies included in the current systematic review, the majority used language sample analysis—a

method that requires a fair amount of expertise--to identify, assess, and compare various aspects of language between groups with typical and atypical language abilities. In addition, some studies determined the presence of LI in their samples using dynamic assessment procedures which involved collecting teacher/caregiver observations in concert with standardized language measures. Per ASHAs guidelines, using nonstandardized language measures and caregiver interviews in conjunction with standardized language assessments is a form of dynamic assessment (ASHA, 2003), which has been identified as a reliable way to identify impairment in children with linguistic differences. Although standardized test scores are generally needed to qualify children for services, the literature is replete with studies citing concerns about the reliability and validity of standardized instruments when used with dialect speakers and ASHA maintains that clinicians have the autonomy to employ dynamic methods to arrive at a clinical diagnosis yet they are seldom used with AAE-speaking children. In fact, in the extant literature, dynamic assessment has been examined mainly with English Language Learners (ELLs) using a test-teach-retest paradigm where it is expected that children with typical language ability those with typical language ability will demonstrate greater positive change in targeted language areas employing cognitive skills (e.g., attention, self-regulation, problem solving, and flexibility) (Gutierrez-Clellen & Peña, 2001; Peña, Gilliam, & Bedore, 2014; Peña, Quinn, & Iglesias, 1992; Peña, Resendiz, & Gillam 2007). Along similar lines, Oetting et al. (2016) stressed the importance of understanding how to use diagnostic indices (i.e., sensitivity and specificity) to select appropriate assessments and collecting family histories to increase diagnostic accuracy.

To ensure that clinicians are competent with implementing dynamic assessment procedures will require providing culturally relevant curricula (Horton-Ikard, Munoz, Thomas-Tate, & Keller-Bell, 2009), challenging the status quo that has historically consisted of "tacking on" one

or two slides to address cultural and linguistic diversity across core subjects, and a conscious effort to diversify university faculty and the student body (Horton-Ikard & Munoz, 2010). Culturally relevant curricula are essential to implementing evidence-based practice. Further, improving service delivery for children from minority backgrounds is largely dependent upon how clinicians are trained, and the research base that they have to support their practice.

In conclusion, the findings of the 19 studies included in the current literature review indicate that various aspects of language have the potential to be clinically informative in language assessments with children who speak AAE. These studies represent a solid beginning to improving assessment methods for children from diverse backgrounds but there is much more work that needs to be done. Improving assessment methods will ensure that children who need services get them as recent research has suggested that language-minority children are underrepresented on special education caseloads, and specifically in the categories of learning disabilities and speech-language impairments (Morgan et al., 2015). Future research focused on increasing our knowledge of AA children's language longitudinally and developing assessment and screening instruments that are culturally and linguistically sensitive is imperative for improving service delivery for AA children from low SES backgrounds. Finally, the findings and research recommendations of this review, although focused on AA children, have implications for other non-mainstream dialects especially for languages in which dialect significantly impacts the mainstream language and are predominantly spoken by individuals from lower social class.

References

American Speech-Language-Hearing Association. (1983). *Social dialects* [Position Statement]. Retrieved from www.asha.org/policy.

American Speech-Language-Hearing Association. (2003). *American English dialects* [Technical Report]. Retrieved from www.asha.org/policy.

Apel, K., & Thomas-Tate, S. (2009). Morphological awareness skills of fourth-grade African American students. *Language, Speech, and Hearing Services in Schools, 40,* 312-324. doi:10.1044/0161-1461(2009/08-0015)

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

*Burns, F. A., de Villiers, P. A., Pearson, B. Z., Champion, T. B. (2012). Dialect-neutral indices of narrative and evaluation. *Language, Speech, and Hearing Services in Schools, 43,* 132-152. doi:10.1044/0161-1461(2011/10-0101)

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (1999). Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. *Scientific Studies of Reading, 3*(4), 331-361. doi: 10.1207/s1532799xssr03042

Champion, T. B., & Hyter, Y. D., McCabe, A., & Bland-Stewart, L. M. (2003). A matter of vocabulary: Performance of low-income African American head start children on the Peabody Picture Vocabulary Test-III. *Communication Disorders Quarterly, 24(1),* 121-127.

Charity, A. H., Scarborough, H. S., & Griffin, D. M. (2004). Familiarity with school English in African American children and its relation to early reading achievement. *Child Development, 75(5),* 1340-1356.

*Cleveland, L. & Oetting, J. (2013). Verbal –s marking by dialect and clinical status. *American Journal of Speech-Language Pathology, 22*(4), 604-614. doi:10.1044/1058-0360(2013/12-0122)

Cole, P. A., & Taylor, O. L. (1990). Performance of working class African American children on three tests of articulation. *Language, Speech, and Hearing Services in Schools, 21,* 171-176. doi:10.1044/0161-1461.2103.171

Craig, H. K., Kolenic, G. E., & Hensel, S. L. (2014). African American English-speaking students: A longitudinal examination of style shifting from kindergarten through second grade. *Journal of Speech, Language, and Hearing Research, 57,* 143-157. doi:10.1044/1092-4388(2013/12-0157)

Craig, H. K., Thompson, C. A., Washington, J. A., & Potter, S. L. (2003). Phonological features of child African American English. *Journal of Speech, Language, and Hearing Research, 46,* 623-635. doi:10.1044/1092-4388(2003/049)

Craig, H. K., Thompson, C. A., Washington, J. A., & Potter, S. L. (2004). Performance of elementary-grade African American students on the Gray Oral Reading Tests. *Language, Speech, and Hearing Services in Schools, 35,* 141–154. doi:10.1044/0161-1461(2004/015)

Craig, H. K., & Washington, J. A. (1994). Dialectal forms during discourse of poor, urban, African American preschoolers, *Journal of Speech and Hearing Research, 37,* 816-823.

*Craig, H. K., & Washington, J. A. (2000). An assessment battery for identifying language impairments in African American children. *Journal of Speech, Language, and Hearing Research, 43,* 366-379. doi:10.1044/jslhr.4302.36

Craig, H. K., & Washington, J. A. (2002). Oral language expectations for African American preschoolers and kindergartners. *American Journal of Speech-Language Pathology*, *11*, 59-70. doi:10.1044/1058-0360(2002/007)

Craig, H. K., & Washington, J. A. (2004). Grade-related changes in the production of African American English. *Journal of Speech, Language, and Hearing Research*, *47*, 450–463. doi:10.1044/1092-4388(2004/036)

Craig, H. K., Washington, J. A., Thompson-Porter, C. (1998). Average c-unit lengths in the discourse of African American children from low-income, urban homes. *Journal of Speech, Language, and Hearing Research*, *41*, 433-444. doi:10.1044/jslhr.4102.43

*de Villiers, J., Roeper, T., Bland-Stewart, L., & Pearson, B. (2008). Answering hard questions: Wh-movement across dialects and disorder. *Applied Psycholinguistics*, *29*, 67-103. doi: 10.1017/S0142716408080041

Dollaghan, C. A., Campbell, T. F., Paradise, J. L., Feldman, H. M., Janosky, J. E., Pitcairn, D. N.. & Kurs-Lasky, M. (1999). Maternal education and measures of early speech and language. *Journal of Speech, Language, and Hearing Research*, *42*, 1432-1443. doi:10.1044/jslhr.4206.1432

Fasold, R. (1969). Tense and the form be in Black English. *Language* 45: 763–76.

Fiorentino, L., & Howe, N. (2004). Language competence, narrative ability, and school readiness in low-income preschool children. *Canadian Journal of Behavioural Science*, *36*(4), 280-294.

Fortier, A. (1884). The French language in Louisiana and the Negro-French dialect. *Transaction of the Modern Language Association of American, 1*, 96-111.

*Garrity, A., & Oetting, J. (2010). Auxiliary BE production by African American English-

speaking children with and without specific language impairmentf. *Journal of Speech, Language, and Hearing Research, 53,* 1307-1320. doi: 10.1044/1092-4388(2010/09-0016)

Gatlin, B., & Wanzek, J. (2015). Relations among children's use of dialect and literacy skills: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 58,* 1306-1318. doi:10.1044/2015_JSLHR-L-14-0311

Gutierrez-Clellen, V. F., & Peña, E. (2001). Dynamic assessment of diverse children: A tutorial. *Language, Speech, and Hearing Services in the Schools, 32,* 212-224. doi:10.1044/0161-1461(2001/019)

Hammer, C. S., Pennock-Roman, M., Rzasa, S., & Tomblin, J. B. (2002). An analysis of the Test of Language Development-Primary for item bias. *American Journal of Speech-language Pathology, 11,* 274-284. doi:10.1044/1058-0360(2002/032)

Hart, B., & Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children.* Baltimore, MD: Paul H. Brookes Publishing Company.

Horton-Ikard, & Apel, K. (2014). Examining the use of spoken dialect indices with African American children in the southern United States. *American Journal of Speech-Language Pathology, 23,* 448-460. doi: 10.1044/2014_AJSLP-13-0028.

Horton-Ikard, R., & Miller, J. F. (2004). It is not just poor kids: the use of AAE forms by African-American school-aged children from middle SES communities. *Journal of Communication Disorders, 37,* 467-487. doi: 10.1016/j.jcomdis.2004.02.001

Horton-Ikard, R., & Munoz, M. L. (2010). Addressing multicultural issues in communication sciences and disorders. *Contemporary Issues in Communication Science and Disorders, 37,* 167-173.

Horton-Ikard, R., Munoz, M. L., Thomas-Tate, S., & Keller-Bell, Y. (2009). Establishing a pedagogical framework for the multicultural course in communication sciences and disorders. *American Journal of Speech-Language Pathology, 18,* 192-206. doi:10.1044/1058-0360(2008/07-0086)

Horton-Ikard, R., Weismer, S. E., & Edwards, C. (2005). Examining the use of standard language production measures in the language samples of African American toddlers. *Journal of Multilingual Communication Disorders, 3*(3), 169-182

Johnson, V. (2005). Comprehension of third person singular in African American children. *Speech, Language, and Hearing in the Schools, 36,* 116-124. doi:10.1044/0161-1461(2005/011)

*Johnson, V. E., & de Villiers, J. G. (2009). Syntactic frames in fast mapping verbs: Effect of age, dialect, and clinical status. *Journal of Speech, Language, and Hearing Research, 52,* 610-622. doi:10.1044/1092-4388(2008/07-0135)

Kresheck, J. D., & Nicolosi, L. (1973). A comparison of Black and White children's scores on the Peabody Picture Vocabulary Test. *Language, Speech, and Hearing Services in Schools, 4,* 37-40. doi:10.1044/0161-1461.0401.37

Labov, W., Cohen, P., Robins, C., & Lewis, J. (1968). A Study of the non-Standard English of Negro and Puerto Rican Speakers in New York City. *Cooperative Research Project, 1,* No. 3288. New York: Columbia University.

*Laing, S. P. (2003). Assessment of phonology in preschool African American Vernacular English speakers using an alternate response mode. *American Journal of Speech-Language Pathology, 12,* 273-281. doi:10.1044/1058-0360(2003/073)

Lau v. Nichols, 414 U.S. 563 (1974).

Leonard, L. B., (1995). Functional Categories in the Grammars of Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research, 38,* 1270-1283. doi:0.1044/jshr.3806.1270

Martin Luther King Junior Elementary School Children et al. v. Ann Arbor School District, 73 F. Supp. 1371 (E.D. Mich. 1979).

Morgan, P. L., Farkas, G., Hillemeier, M. M., Mattison, R., Maczuga, S., Li, H., & Cook, M. (2015). Minorities are disproportionately underrepresented in special education Longitudinal evidence across five disability categories. *Educational Researcher, 44*(5), 278-292. doi:10.3102/0013189X15591157.

Morgan, P. L., Hammer, C. S., Farkas, G., Hillemeier, M. M., Maczuga, S., Cook, M., & Morano, S. (2016). Who receives speech/language services by 5 years of age in the United States? *American Journal of Speech-Language Pathology, 25*, 183-199. doi:10.1044/2015_AJSLP-14-0201

Newkirk-Turner, B. L., Oetting, J. B., & Stockman, I. J. (2014). BE, DO and modal auxiliaries or 3-year-old African American English Speakers. *Journal of Speech, Language, and Hearing Research, 57*, 1383-1393. doi:10.1044/2014_JSLHR-L-13-0063

*Oetting, J. B., & Cleveland, L.H. (2006). The clinical utility of nonword repetition for children living in the rural south of the US. Clinical Linguistics and Phonetics, 20, 553-561. doi: 10.1080/02699200500266455

Oetting, J. B., Gregory, K. D., & Riviere, A. M. (2016a). Changing how speech-language pathologists think and talk about dialect variation. *Perspectives on the ASHA Special Interest Group 16, 1,* 2837. doi:10.1044/persp1.SIG16.28

*Oetting, J., & McDonald, J. (2001). Nonmainstream dialect use and specific language impairment. *Journal of Speech, Language, and Hearing Research, 44,* 207-223. doi:10.1044/1092-4388(2001/018)

Oetting, J., & McDonald, J. (2002). Methods for characterizing participants' nonmainstream dialect use in child language research. *Journal of Speech, Language, and Hearing Research, 45*, 505-518. doi:10.1044/1092-4388(2002/040)

*Oetting, J. B., McDonald, J. L., Seidel, C. M., & Hegarty, M. (2016b). Sentence recall by children with SLI across two nonmainstream dialects of English. *Journal of Speech, Language, Hearing Research, 59,* 183-194. doi:10.1044/2015_JSLHR-L-15-0036

*Oetting, J., & Newkirk, B. (2008). Subject relative clause use by children with and without SLI across dialects. *Clinical Linguistics and Phonetics, 22*, 111-125.

*Oetting, J. B., Newkirk, B. L., Hartfield, L. R., Wynn, C. G., Pruitt, S. L., & Garrity, A. W. (2010). Index of productive syntax for children who speak African American English. *Language, Speech, and Hearing Services in Schools, 41,* 328-339. doi:10.1044/0161-1461(2009/08-0077)

Oetting, J. B., Riviere, A., Lee, R., Richardson, J., Hegarty, M., & McDonald, J. (2012). *Structure Specific Differences in Nonmainstream Dialects of English Across Tasks* . Poster session presented at the Symposium on Research Child Language Disorders, Madison, WI.

Pankratz, M. E., Plante, E., Vance, R., & Insalaco, D. M. (2007). The diagnostic and predictive validity of the Renfrew Bus Story. *Language, Speech, and Hearing Services in Schools, 38*, 390-399. doi:10.1044/0161-1461(2007/040)

Pearson, B. Z., Jackson, J. E., Wu, H. (2014). Seeking a valid gold standard for an innovative, dialect-neutral language test. *Journal of Speech, Language, and Hearing Research, 57,* 495-508. doi: 10.1044/2013_JSLHR-L-12-0126

Peña, E. D., Gillam, R. B., & Bedore, L. M. (2014). Dynamic assessment of narrative ability in English accurately identifies language impairment in English language learners. *Journal of Speech, Language, and Hearing Research, 57,* 2208-2220. doi:10.1044/2014_JSLHR-L-13-0151

Peña, E., Quinn, R. & Iglesias, A. (1992). The application of dynamic methods to language assessment: A nonbiased procedure. *The Journal of Special Education, 26*(3), 269-280.

Peña, E. D., Resendiz, M. & Gillam, R. B. (2007). The role of clinical judgments of modifiability in the diagnosis of language impairment. *Advances in Speech-Language Pathology, 9*(4), 332-345.

Plante, E., Swisher, L., Kiernan, B., & Restrepo, A. (1993). Language matches: Illuminating or confounding. *Journal of Speech and Hearing Research, 36,* 772-776. doi:10.1044/jshr.3604.772

Pruitt, S., & Oetting, J. (2009). Past tense marking by African American English-speaking children reared in poverty. *Journal of Speech, Language, and Hearing Research, 52,* 2-15. doi:10.1044/1092-4388(2008/07-0176)

Restrepo, M. A., Schwanenflugel, P. J., Blake, J., Neuharth-Pritchett, S., Cramer, E. S., & Ruston, H. P. (2006). Performance on the PPVT-III and the EVT: Applicability of the measures with African American and European American preschool children. *Language, Speech, and Hearing Services in Schools, 37,* 17-27. doi:10.1044/0161-1461(2006/003)

Rickford, J. (1999). *African American Vernacular English.* Malden, MA: Blackwell Publishers Inc.

Rickford, J., & Rafal, C. (1996). Preterite Had + V-Ed in the narratives of African-American preadolescents. *American Speech, 71,* 227-254.

*Rodekohr, R. K., & Haynes, W. O. (2001). Differentiating dialect from disorder: A comparison of two processing tasks and a standardized language test. *Journal of Communication Disorders, 34,* 255-272.

Ross, S., Oetting, J., & Stapleton, B. (2004). Preterite HAD + -ED: A developmental narrative structure of African American English. *American Speech, 79.*

Roy, J., Oetting, J., Moland, C. (2013). Linguistic constraints on children's overt marking of BE by dialect and age. *Journal of Speech, Language, and Hearing Research, 56,* 933-944. doi:10.1044/1092-4388(2012/12-0099)

Seymour, H. N. (2004). The challenge of language assessment for African American English-speaking children: A historical perspective. *Research on the acquisition of AAE by children:* Paper 1. Retrieved from http://scholarworks.umass.edu/csaal_child/1.

*Seymour, H., Bland-Steward, L., & Green, L. (1998). Difference versus deficit in child African American English. *Language, Speech, and Hearing Services in the Schools, 29*, 113-138.

Seymour, H., Roeper, T. W., & de Villiers, J. (2003). *Diagnostic Evaluation of Language Variation—Screening Test*. San Antonio, TX: Harcourt Assessment.

Seymour, H. N., Roeper, T. W., de Villiers, J. (2005). *Diagnostic Evaluation of Language Variation-Norm Referenced.* Minneapolis, MN: Pearson.

Stockman, I. J. (1996). The promises and pitfalls of language sample analysis as an assessment tool for linguistic minority children. *Language, Speech, and Hearing Services in Schools, 27,* 355-366.

Stockman, I. J. (2006). Alveolar bias in final consonant deletion patterns of African American children. *Language, Speech, and Hearing Services in Schools, 37,* 85-95.

*Stockman, I. J. (2008). Toward validation of a minimal competence phonetic core for African American children. *Journal of Speech, Language, Hearing Research, 51,* 1244-1262. doi:10.1044/1092-4388(2008/07-0081)

Stockman, I. J. (2010). A review of developmental and applied language research on African American children: From a deficit to a difference perspective on dialect differences. *Language, Speech, and Hearing Services in Schools, 41,* 23-38. doi:10.1044/0161-1461(2009/08-0086)

*Stockman, I. J., Guillory, B., Seibert, M., & Boult, J. (2013). Toward validation of a minimal competence core of morphosyntax for African American children. *American Journal of Speech-Language Pathology, 22,* 40-56. doi: 10.1044/1058-0360(2012/11-0124)

*Stockman, I. J., Newkirk-Turner, B. L., Swartzlander, E., & Morris, L. R. (2016) Comparison of African American children's performances on a minimal competence core for morphosyntax and the Index of Productive Syntax. *American Journal of Speech-Language Pathology, 25,* 80-96. doi: 10.1044/2015_AJSLP-14-0207

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934-947.

Terry, N. P. (2006). Relations between dialect variation, grammar, and early spelling skills. *Reading and Writing, 19,* 907-931. doi: 10.1007/s11145-006-9023-0

Terry, N. P. (2012). Dialect variation and phonological knowledge: Phonological representations and metalinguistic awareness among beginning readers who speak nonmainstream American English. *Applied Psycholinguistics,* 1-22. doi:10.1017/S0142716412000276

Terry, N., Connor, C., Thomas-Tate, S., & Love, M. (2010). Examining relationships among dialect variation, literacy skills, and school context in first grade. *Journal of Speech, Language, and Hearing Research, 53,* 126-145. doi:10.1044/1092-4388(2009/08-0058)

Terry, N. P., & Scarborough, H. (2011). The phonological hypothesis as a valuable framework for studying the relation of dialect variation to early reading skills. In S. Brady, D. Braze, & C. A. Fowler (Eds.), Explaining individual differences in reading: Theory and evidence (pp. 97–117).

Thomas-Tate, S. R., Washington, J. A., Craig, H. K., & Packard, M. E. (2006). Performance of African American preschool and kindergarten students on the Expressive Vocabulary Test. *Language, Speech, and Hearing Services in Schools*, 37, 143–149. doi:10.1044/0161-1461(2006/016)

Thomas-Tate, S., Washington, J., & Edwards, J. (2004). Standardized assessment of phonological awareness skills in low-income African American first graders. *American Journal of Speech-Language Pathology, 13,* 182–190. doi:10.1044/1058-0360(2004/018)

Thompson, C. A., Craig, H. K., & Washington, J. A. (2004). Variable production of African American English across oracy and literacy context. *Language, Speech, and Hearing Services in Schools, 35,* 269-282. doi:10.1044/0161-1461(2004/025)

Washington, J. A., & Craig, H. K. (1992). Performances of low-income, African American preschool and kindergarten children on the Peabody Picture Vocabulary Test-Revised. *Language, Speech, and Hearing Services in Schools, 23,* 329-333. doi:10.1044/0161-1461.2304.32

Washington, J., & Craig, H. (1994). Dialectal forms during discourse of poor, urban, African American Preschoolers. *Journal of Speech and Hearing Research, 37*, 816-823. doi:10.1044/jshr.3704.816

Washington, J. A., & Craig, H. K. (1998). Socioeconomic status and gender influences on children's dialectal patterns. *Journal of Speech, Language, and Hearing Research, 41*, 618-626. doi:10.1044/jslhr.4103.618

Washington, J. A., & Craig, H. K. (1999). Performance of at-risk, African American preschoolers on the Peabody Picture Vocabulary Test-III. *Language, Speech, and Hearing Services in Schools, 30,* 75-82. doi:10.1044/0161-1461.3001.7

*Washington, J. A., & Craig, H. K. (2004). A language screening protocol for use with young African American children in urban settings. *American Journal of Speech-Language Pathology, 13*, 329-340. doi:10.1044/1058-0360(2004/033)

Washington, J. A., Craig, H. K., Kushmaul, A. J. (1998). Variable use of African American English across two language sampling contexts *Journal of Speech, Language, and Hearing Research, 41*, 1115-1124. doi:10.1044/jslhr.4105.1115

Weismer, S. E., Murray-Branch, J., & Miller, J. F. (1994). A prospective longitudinal study of language development in late talkers. *Journal of Speech, Language, and Hearing Research, 37*, 852-867. doi:10.1044/jshr.3704.852

*Wilcox, L. D., & Anderson, R. T. (1998). Distinguishing between phonological difference and disorder in children who speak African-American Vernacular English: An experimental testing instrument. *Journal of Communication Disorders, 31,* 315-335.

Wolfram, W. (1969). *A Sociolinguistic Description of Detroit Negro Speech.* Washington, D.C.: Center for Applied Linguistics.

Wolfram, W. (1998). Language ideology and dialect: Understanding the Oakland Ebonics controversy. *Journal of English Linguistics, 26*(2), 108-121.

Wolfram, W., & Fasold, R. (1974). *The Study of Social Dialects in American English.* Englewood, NJ: Prentice Hall.

Wolfram, W., Temple Adger, C., Christian, D. (1999). *Dialects in schools and communities*. Mahwah, NJ: Lawrence Erlbaum Associates.

Wolfram, W., & Thomas, E. (2002). *The Development of African American English*. Oxford, UK: Blackwell Publishers.

Zimmerman, I. L., Steiner, V. V., Pond, R. E. (2001). *Preschool-Language Scale-Third Edition.* Austin, TX: The Psychological Corporation.

2 GRAMMATICALITY JUDGMENTS OF LOW-INCOME, AFRICAN AMERICAN ENGLISH-SPEAKING CHILDREN: THE ROLE OF LANGUAGE ABILITY AND DIALECT DENSITY

The study of African American English (AAE), a major dialect of American English, has been primarily focused on its expressive language characteristics, specifically morphosyntactic features that contrast with Mainstream American English (MAE). This work has demonstrated that AAE impacts expressive morphology in predictable ways. For example, omission of past tense (e.g., He kickØ the ball.) and copula (e.g., That girl Ø pretty.) and auxiliary (e.g., That boy Ø going to the store.) BE forms (e.g., is, are, was, were) are prominent features of child and adult AAE (Oetting & McDonald, 2001; Washington & Craig, 1994; Wolfram & Thomas, 2002). And whether or not a grammatical morpheme is omitted can be predicted by the surrounding lexical and discourse context. Using past tense to illustrate this point, Rickford (1999) and others (Pruitt & Oetting, 2009) have observed that the regular past morpheme (-ed) is more likely to be omitted when it occurs in a consonant cluster (e.g., kicked) than after a vowel (e.g., played) and children's propensity to omit grammatical features is directly related to their dialect density. Studies have reported that dialect use occurs on a continuum from low to high usage (Horton-Ikard & Miller, 2004; Oetting & McDonald, 2002; Terry, Connor, Thomas-Tate, & Love, 2010; Washington & Craig, 1994), and it is now well understood that a child's place on this continuum (i.e., low, moderate, high) influences performance on a variety of language and literacy measures (Charity, Scarborough, & Griffin, 2004; Connor & Craig, 2006; Craig, Thompson, Washington, & Potter, 2004; Craig & Washington, 2004; Gatlin & Wanzek, 2016; Terry, 2006; Terry et al., 2010)

The impact of AAE use on receptive language is less clear. Some early studies examining this relationship have demonstrated that language and reading comprehension are impacted by nonstandard dialect use (Nelson & McCrosky, 1978; Steffensen, Reynolds, McClure, & Guthrie, 1981) while others have observed no significant impact (Gemake, 1981; Johnson, & Simons, 1973). More recent examples of these mixed findings include and Craig and colleagues (2004), in which reading rate and accuracy were predicted by dialect use but there was no association between dialect use and reading comprehension. In contrast, Charity and colleagues (2004) reported that dialect use was negatively associated with measures of reading accuracy and reading comprehension.

Unfortunately, since receptive language has not received systematic, sustained attention in the extant literature and existing findings are inconclusive, it remains unclear if, or how, dialect use impacts development of receptive language knowledge. This is a critical omission in the research base since receptive and expressive aspects of language are essential to developing a complete and comprehensive profile of any child's language abilities (Tomblin, Records, & Zhang, 1996). Though more recent studies show that being a dialect speaker may impact comprehension of certain morphosyntactic structures that are contrastive with MAE (de Villiers & Johnson, 2007; Johnson, 2005), the relationship between AAE use and receptive language ability is worth clarifying in the interest of improving service delivery for African American (AA) children, particularly those growing up in poverty as these children tend to be high dialect users (Craig & Washington, 2004; Washington & Craig, 1998).

The purpose of this investigation was to examine receptive language abilities in a sample of low-income, AAE-speaking children whose dialect density and language ability differed along a continuum, using grammaticality judgments. Grammaticality judgments are an important in-

dex of receptive grammatical knowledge, providing a measure of children's morphosyntactic awareness/knowledge. Children's grammatical judgments will be examined relative to their general language ability and their dialect density, as this has been informative for characterizing language used by AA children in both oral and written contexts (Oetting & McDonald, 2001; Pruitt & Oetting, 2009; Thompson, Craig, & Washington, 2004; Washington, Craig, Kushmaul, 1998). Examining AAE-speaking children's grammaticality judgments should contribute to the development of a more comprehensive language profile of AAE, leading to improved diagnostic accuracy for this population.

Grammaticality Judgment

The ability to make accurate grammatical judgments reflects an individual's linguistic competence. In grammaticality judgment contexts children are asked to indicate whether sentences are grammatical or ungrammatical/correct or incorrect (Wulfeck, 1993). In MAE, grammaticality judgment ability has been used as a clinical indicator of impairment as it has been demonstrated that children with low language skills struggle to make accurate judgments (Rice, Wexler, & Redmond, 1999). There is a paucity of grammaticality judgment research studies focused on AAE speakers. However, findings from reading research using *phonological* judgments (Terry, 2012; Terry & Scarborough, 2011) suggest this paradigm may be useful for accessing linguistic knowledge of AAE speakers.

Grammaticality Judgment in MAE/LI

Studies focused on MAE speakers show that children with language impairment (LI), from early childhood through adolescence, are less accurate in their judgments than their peers who are typically developing (Miller, Leonard, & Finneran, 2008; Noonan, Redmond, & Archibald, 2014; Pawlowska, Robinson, & Seddoh, 2013; Redmond & Rice, 2001; Rice, Hoffman, &

Wexler, 2009; Rice et al., 1999; Wulfeck, Bates, Krupa-Kwiatkowski, & Saltzman, 2004). These studies, using a grammaticality judgment paradigm, report that, in particular, children with LI have marked difficulty detecting tense-marking errors. Rice and colleagues (1999) examined the grammatical judgment abilities of six-year-old, MAE-speaking children with and without specific language impairment (SLI) in a longitudinal study. Examiners acted out a conversation between two toy robots (in addition to using other relevant props) and participants were asked to verbally indicate whether sentences spoken by the robots were "good" or "not so good". Children with SLI showed reduced ability to detect errors of grammatical omissions involving subject-verb agreement with BE (e.g., *he Ø running away*) and third-person singular (e.g., *he eatØ toast*), when compared to typical controls.

In a similar vein, Pawlowska and colleagues (2014) used a judgment task to demonstrate differences in judgment ability based on various aspects of language. Their findings demonstrated that while tense marking is difficult for children with LI to judge, they have less difficulty making judgments about non-tense and non-agreement structures. For example, four- and five-year-olds with and without LI in this study identified lexical errors at higher rates than morpho-syntactic errors although significant group differences were found for both language domains favoring the typical group. Wulfeck and colleagues (2004) also reported that although overall grammaticality judgment ability of the SLI group lagged behind the typical controls, children's judgments of word order errors improved with age at a higher rate than children's ability to judge agreement errors. Non-tense and non-agreement related items are also easier for younger children who are typically developing. For example, McDonald (2008) reported that typical children ages six to 11-years-old performed comparably when judging word order errors but noted developmental differences favoring the older children on tense and subject-verb agreement items.

Taken together these studies demonstrate that judgment tasks, in particular those that assess knowledge of grammatical structures, are very informative for use with LI children who speak MAE and for children whose language varies from MAE. These tasks successfully distinguished children who had LI from their typical language peers. .

Grammaticality Judgment in AAE.

There are few published grammaticality judgment studies focused on the oral language of AAE speakers. There are, however, studies that use the judgment paradigm to examine *phonological* judgments of AAE speakers; these studies are focused on reading abilities (Terry, 2012; Terry & Scarborough, 2011). Both of these investigations examined children's knowledge of word pronunciations as a function of dialect density. Stimuli reflected MAE pronunciations (e.g., *breakfast*), pronunciations consistent with AAE (e.g., *breffis*), and those that were not specific to either dialect (e.g., *bweakfast*). Results showed low and high dialect users judged MAE pronunciations at equivalent rates, though high dialect users were more accepting of both AAE and non-dialect specific pronunciations than were children in the who were moderate AAE users.

In a preliminary study that was specifically focused on morphosyntactic grammaticality judgments, Lee James, Washington, and Seidenberg (submitted) examined the judgments of 179 AAE-speaking children ages eight- to 10-years-old who had been identified as either low or high dialect speakers. Data were collected using a standardized grammaticality judgment assessment. Analyses routine involved comparing means routines and revealed children's judgments were influenced by dialect density and item type. More specifically, children differed in their responses to AAE items based on their dialect density: the high dialect group was more likely to judge AAE items as correct than the low dialect group. Further, children evidenced similarities

in their responses to MAE and ungrammatical items (not consistent with either AAE or MAE grammar). Both dialect groups were highly accurate in their judgments of MAE items and uncertain in their judgments of ungrammatical sentences resulting in chance level performance on some items.

The results of James, Washington, and Seidenberg, though preliminary, support the findings of Terry and Scarborough (2011) and Terry (2012) which found that: 1) young dialect speakers were likely to judge MAE and AAE pronunciations as correct, showing knowledge of both MAE and AAE word structure; and, 2) children's tendency to accept or reject AAE pronunciations was related to their dialect density. Taken together these studies show that judgment tasks can be effective in gleaning information regarding children's receptive phonological and morphological knowledge. The current study will extend this work by including a description of grammatical knowledge in a larger sample of dialect speakers whose language skills vary from below to above average**.**

Present Study

The purpose of the current study is to characterize the grammaticality judgments of second through fifth grade, low-income AAE speakers whose language ability and dialect density differ along a continuum. Studies have evidenced that the grammaticality judgment paradigm is clinically useful for distinguishing typical children from children with LI among school-aged MAE speakers. It is important to assess its usefulness with AAE speakers whose language use includes many differences from MAE in the use of grammatical morphemes, including variable inclusion and exclusion. In addition, it has been documented that children growing up in poverty use more nonmainstream features than do their middle-income peers (Craig & Washington, 2004; Washington & Craig, 1998), contributing to the difficulty of identifying LI in this popula-

tion. Therefore, to avoid the potential confound of socioeconomic status, the current study focuses solely on participants from impoverished backgrounds. This study aims to contribute to the developmental profile of AAE speakers growing up in poverty by providing a description of children's receptive morphological knowledge, as indexed by their grammaticality judgments. Previous research on MAE speakers indicates that judgment ability is impacted by age, language ability, and item type. In their sample of low-income AAE speakers, James and colleagues (submitted) found grammaticality judgments were impacted by degree of dialect use—a finding that reflects the morphosyntactic nature of the dialect. The current study will explore the contribution of language ability, and dialect density to children's performance on a standardized grammaticality judgment measure as a preliminary step in evaluating the clinical utility of this task for AAE speakers. Intuitively it makes sense that children with typical language ability will outperform children with low language skills on the grammaticality judgment task thus it is expected that children's grammaticality judgments will be impacted by their dialect density as well as their language ability.

The morphological comprehension subtest of the Test of Language Development-Intermediate: Fourth Edition (TOLD-I:4; Hammill & Newcomer, 2008) will be used to assess children's grammaticality judgment ability. The following research questions are posed:

1. What is the overall relationship between dialect density and language ability in a sample of low-income AAE speakers?
2. When considering performance on a grammaticality judgment task in a sample of low-income AAE speakers:
    a. What is the contribution of dialect density and language ability on overall performance as measured by standard scaled scores?

b. Is there an interaction between language ability and dialect density on children's overall performance?

c. What is the contribution of dialect density and language ability on specific item types (i.e., AAE, MAE and ungrammatical)?

Methodology

Participants

Participants in this investigation were 273, AA children enrolled in second through fifth grades, ranging in age from 8;0 to 11;11 years old. One hundred seventy nine children in the current sample also participated in the James, Washington, and Seidenberg (submitted) study. All participants were from low-income backgrounds and had nonverbal intelligence within normal limits (M = 102.52, SD = 9.91) on the Kaufman Brief Intelligence Test: Second Edition (KBIT-2; Kaufman & Kaufman, 2004). Free and reduced priced lunch (FRPL) was used as an index of income status and all participants attended schools where 87% to 100% of the student body was eligible for FRPL. All participants were enrolled in a larger project focused on language and literacy outcomes for urban AA students that included assessment of language, reading, cognitive and writing skills. Data for the project were collected by trained master's level and doctoral level students in communication sciences and disorders and other related fields (e.g., counseling, early childhood education, psychology). Examiners were from a variety of racial and ethnic backgrounds.

Dialect status was confirmed and quantified using the Diagnostic Evaluation of Language Variation-Screening Test (DELV-ST; Seymour, Roeper, & de Villiers, 2003). Based on the participant's response pattern, the screener quantified children's dialect use on a continuum using three classifications including: *MAE, some variation from MAE*, and *strong variation from*

MAE. Majority of the participants (68.5%) were classified as presenting with *strong* variation from MAE. Data from the current study support claims in the literature that associate nonmainstream dialect use with lower social class (Antoniou, Grohmann, Kambanaros, & Katsos, 2016) and greater dialect use among lower income communities compared with individuals from higher income backgrounds (Washington & Craig, 1998).

Participants were also classified according to language ability. A battery of subtests from the two standardized language assessments, CTOPP-2 (Wagner, Torgesen, Rashotte, & Pearson, 2013) and the TOLD-I:4 were administered to each participant. Subtests administered from the CTOPP-2 included Elision, Blending Words, Memory for Digits, Nonword Repetition, and Rapid Letter Naming. Subtests administered from the TOLD-I:4 included Picture Vocabulary and Sentence Combining. Each subtest was administered and scored in accordance with the procedures in the Examiner's Manuals. Because individual subtests were administered and not the entire assessment it was not possible to calculate an overall scaled score for each measure. Instead, children were grouped based on an aggregate mean scaled score from the CTOPP-2 and TOLD-I:4. To compute an aggregate mean, first, an average scaled score was computed by aggregating the subtests for each measure resulting in a mean scaled score for each participant on the CTOPP-2 and the TOLD-I:4. Then, a composite scaled score was computed by averaging each participant's mean scaled scores from the CTOPP-2 and TOLD-I:4. The mean and standard deviation for the language composite were 8.08 and 1.66, respectively. The Shapiro Wilkes Test of Normality yielded a p-value of .032 which indicated that the data for the current sample were not normally distributed however skewness (.147) and kurtosis (-.119) values indicated a symmetrical distribution. It is important to note that the computed mean for the language composite (M = 8.08, SD = 1.66, range 4-13) is approximately two scaled points below the published mean

of 10 (*SD* = 3) for the CTOPP-2 and TOLD-I:4 revealing that average performance in the current sample is indeed below average when compared to the normative sample.

Children were classified into language ability groups relative to the distribution of the current sample. The composite variable was used to classify participants into the following groups: below average, average, and above average. Children who scored more than -1 *SD* (< 6.42) of the mean were identified as having below average language skills (*n* = 47), children who earned scaled scores between –1 *SD* and +1 *SD* (6.42 - 9.74) were identified as having average language skills (*n* = 182), and children who earned scores greater than +1 *SD* (> 9.74) of the mean were identified as above average language skills (*n* = 44). Participant data is displayed in Table 1.

Table 1

Participant Data by Dialect and Language Ability

	Age (in years)				Gender			
Dialect Density	8	9	10	11	M	F	KBIT SS *M* (*SD*)	
Low	40	25	19	9	41	52	104.62	(10.43)
High	53	54	54	19	89	90	101.44	(9.48)
Total	93	79	73	28	130	142	103.03	(9.55)
Language Ability								
Below	8	13	19	7	25	22	98.23	(8.89)
Average	65	52	46	19	87	94	102.28	(9.62)
Above	20	14	8	2	18	26	108.11	(9.72)
Total	93	79	73	28	130	142	102.87	(9.41)

Note. The KBIT has a $M = 100$, $SD = 15$. KBIT = Kaufman Brief Intelligence Test-Second Edition (KBIT-2; Kaufman & Kaufman, 2004); SS = Standard Scaled Score

Grammaticality Judgment Measure

The Morphological Comprehension (MC) subtest from the TOLD-I:4 was used to examine children's grammaticality judgments ability. The TOLD-I:4 is normed on children ages 8;0 to 17;11 and encompasses subtest that measure language skills related to listening, organizing, and speaking. The MC subtest was designed to measure children's grammatical knowledge through their responses to grammatical and ungrammatical sentences. Children are asked to indicate whether the sentence spoken by the examiner sounds "correct" or "incorrect". The premise is that children with low language ability will lack the underlying morphosyntactic knowledge needed to determine a sentence's grammaticality. Evidence from MAE-speaking

children indicates that children with SLI are significantly less accurate in their ability to make grammatical judgments when compared to their typically-developing peers (Rice et al., 1999).

Administration Procedures

Data collectors were trained to administer the MC subtest as it is outlined in the Examiner's Manual. Examiners begin by reading the following directions: "I'm going to say some sentences. Some of the sentences are correct; some are incorrect. You tell me if the sentence I say is correct or incorrect. Let's try one." The TOLD-I:4 is a ceiling test and the procedures indicate that regardless of performance, items 1 - 10 are to be administered unless a participant responds incorrectly to two foils. The ceiling rule requires the examiner end testing is if the participant misses three out of five items beginning with item 11. The study conducted by James, Washington, and Seidenberg (under review) found that a significant portion of the sample reached a ceiling within the first 15 items. This was partially attributed to the fact that dialect-appropriate items are heavily weighted toward the beginning of the test. AAE speakers were more likely to judge these items as correct, so they reached ceiling quickly. Data for the current sample support these findings as at least two thirds of participants reached ceiling before item 16. For this reason, only the first 15 items (including the first 3 foils) were examined in the current study.

Coding and Scoring

The MC subtest was administered and scored according to the procedures outlined in the Examiner's Manual. The measure consists of 56 items total. Six are deemed grammatical sentences consistent with Mainstream grammar (e.g., They like school.) according to the manual. These items were coded as MAE. The remaining 50 are considered ungrammatical in MAE however of these 50 items 25 reflect grammatical sentence structure in AAE including subject verb agreement with BE and DON'T, zero past tense, and alternative pronoun use. These items

were coded as AAE. The remaining 25 ungrammatical items include colloquial structures that are common in speech but not exclusive to either MAE or AAE morphosyntax. These items were coded as ungrammatical (UNG). Items analyzed in the current study included three MAE items, nine AAE items, and 6 UNG items. See Appendix A for items included in the current study.

The scoring for the MC subtest is binary where 1 represents a correct response and 0 represents an incorrect response. To reduce examiner error, raw scores (tallies of correct responses except MAE items) were recorded online at the time of testing and later entered into a software program, which was used to derive scaled scores for each participant. The software computes the participant's age at the time of testing and assigns standard scaled scores accordingly from a predetermined table.

Reliability

Data quality was ensured by conducting entry reliability as well as scoring reliability. This was achieved by double entering and scoring approximately 20% of the sample. Scores were calculated electronically using scoring tables that were created based on raw scores and age equivalents. Agreement for entry reliability was 96% and scoring reliability was 100%.

Results

Relationship between Dialect Density and Language Ability

To address the first research question, "*what is the overall relationship between dialect density and language ability?",* a Point Biserial Correlation was conducted. Dialect density (dichotomous; low, high) and language ability (continuous) were entered into the analysis with gender as a covariate to control for the potential contribution of gender on dialect density (Washington & Craig, 1998). Results showed a significant, negative, moderate correlation between

dialect density and the overall language composite indicating that as dialect density increased from low to high, language ability decreased. The coefficient of determination, r^2 = .17 indicates that dialect density accounted for approximately 17% of the variance in scores between language ability groups. Inspection of the raw data revealed that 100% (47/47) of children with below average language ability were also high dialect users. Importantly, 32% (14/44) of children with above average language ability were also high dialect users. See Table 2 for correlational analysis statistics.

Table 2

Dialect Density and Language Ability: Correlations and Descriptive Statistics

Variables	1	2
1. Dialect Density	-	
2. Language Ability	-.414*	-
M		8.08
SD		1.66
Range		4-13

*p < .001

Grammaticality Judgment: The Role of Dialect Density and Language Ability

Linear multiple regression was used to answer research question 2a, "*what is the contribution of dialect density and language ability on overall grammaticality judgment performance?"*. First, the potential impact of gender on dialect density and overall grammaticality judgment performance was test using correlational analysis. There was no significant association between gender and dialect density (p = .379) nor gender and grammaticality judgment performance (p = .272). Therefore, gender was not considered further.

The final regression model included dialect density and language composite scores as predictors and standard scaled scores from the MC subtest was included as the dependent variable. The initial regression model included standard scores from the KBIT as a predictor because it was expected that nonverbal intelligence might influence children's grammatical knowledge However, findings indicated that performance on the KBIT did not significantly (p = .384) contribute to children's grammaticality judgments. With the KBIT removed, results from the final regression model indicated that dialect density and language ability together explained 33.1% of the variance in grammatical judgment scores. Specifically, dialect density, significantly predicted grammaticality judgment performance overall, as did language ability. See Table 3 for regression data.

Table 3

Summary of Regression Analysis

Variables	B	SE B	β
Dialect Density	-1.31	.279	-.255
Language Ability	.624	.080	.426

Note. adj. R^2 = .331, p < .001

Research question 2b, "*is there an interaction between language ability and dialect density on children's overall performance?",* was addressed by conducting an ANOVA in which dialect density and language ability group were the independent variables. Different from the previous analyses where language ability was a continuous variable (i.e., standard scaled scores), in this analysis language ability was a categorical variable representing three language groups derived from the current sample's distribution of scores—below average (below -1 *SD* or stand-

ard scaled scores < 6.42), average (between -1 *SD* and +1 *SD* or standard scaled scores between 6.42 - 9.74), and above average (above +1 SD or standard scaled scores > 9.74).

Table 4 displays participant means and standard deviations of standard scaled scores by dialect density and language ability group. The main effect for dialect density was statistically significant, $F(1, 268) = 24.08$, $p < .001$, partial $\eta^2 = .082$, revealing higher grammaticality judgment performance for the low dialect group ($M = 8.44$, $SD = 2.63$) compared to the high dialect users ($M = 6.23$, $SD = 1.94$). There was a statistically significant, main effect for language group, $F(2, 268) = 49.94$, $p < .001$, partial $\eta^2 = .080$, and Tukey's post hoc tests yielded significant differences on all comparisons: children with below average language ability ($M = 5.30$, $SD = 1.64$) scored significantly ($p < .001$) lower than children in the average ($M = 6.91$, $SD = 2.17$) and above average (M = 9.09, SD = 2.64) language groups. Children in the average language group also scored significantly lower ($p < .001$) than children with above average language skills. The interaction term was not statistically significant, $F(1, 268) = 1.834$, $p = .177$, partial $\eta^2 = .007$.

Table 4

Overall Performance on the MC Subtest by Dialect Density and Language Ability

Dialect Density	Language Ability	Scaled Scores M (SD)	Range	Total SS for Language Groups
Low (n = 93)	Average (n = 63)	7.78 (2.40)	3-14	
	Above (n = 30)	9.83 (2.58)	5-16	
Total		8.44 (2.63)	3-16	
High (n = 180)	Below (n = 47)	5.30 (1.64)	2-9	5.30 (1.64)
	Average (n = 119)	6.45 (1.90)	3-14	6.91 (2.73)
	Above (n = 14)	7.50 (2.03	5-11	9.09 (2.64)
Total		6.23 (1.93)	2-16	6.99 (2.43)

Finally, a multivariate analysis of variance (MANOVA) was conducted to answer research question 2c, "*How do dialect density and language ability impact performance on specific item types (i.e., AAE, MAE and ungrammatical)?".* Dialect density and language ability groups were the independent variables and three dependent variables, one for each item type, were entered into the analysis. The variables reflected proportion of accurate responses for each item type. For a given item, proportion correct was computed by summing the responses for that item and dividing by either 9 (AAE items), or 3 (MAE items), or 6 (UNG items). The TOLD-I:4 administration procedures requires children to verbally indicate whether the sentence read by the examiner is correct or incorrect and their responses are scored used a in a binary system (1 for an accurate response and 0 for an inaccurate responses). A score of 1 for a given item reflects an accurate response obtained by either, 1) judging AAE and UNG items as incorrect or, 2) judging

MAE items as correct. A score of 0 for a given item reflects an inaccurate response obtained by either, 1) judging AAE and UNG items as correct or, 2) judging MAE items as incorrect.

Table 5 displays the mean proportion accurate and standard deviations for all three item types as a function of both dialect density and language ability groups. Results from the MANOVA overall indicated no statistically significant interaction (p = .588) between dialect density and language ability however the main effect for dialect density, F (3, 266) = 5.72, p < .001, Wilks's Λ = .939, partial η^2 = .061 and the main effect for language ability, F (6, 532) = 4.60, p < .001, Wilk's Λ = .904, partial η^2 = .049, were statistically significant.

With regard to specific item types, there was no statistically significant interaction between dialect density and language ability on any of the item comparisons: AAE (p = .997), MAE (p = .737), and UNG (p = .225). However, the effect for dialect density group was statistically significant for the AAE items only, F (1, 268) = 15.18, p < .001, partial η^2 = .054. This finding reveals that children in the low dialect group identified AAE items as incorrect at a higher proportion compared to high dialect speakers. No statistically significant performance differences were observed between dialect groups on the MAE (p = .137) or UNG items (p = .465).

For language ability groups, statistically significant performance differences were observed on both AAE, F (2, 268) = 9.20, p < .001, partial η^2 = .064, and UNG items, F (2, 268) = 7.69, p = .001, partial η^2 = .054, but not for MAE items (p = .343). Each pairwise comparison for the three language ability groups was statistically significant for AAE and UNG items (p < .05). The post hoc test revealed no significant differences on MAE items among any of the three language ability groups (p > .05). Overall these results reveal that as language ability increases, children's predisposition to judge AAE and UNG items as incorrect increases.

Table 5

Proportion Accurate for Each Item Type by Dialect Density and Language Ability Groups

Dialect Density	Language Ability	AAE *M* (*SD*)	MAE *M* (*SD*)	UNG M (*SD*)
Low (n = 93)	Average (n = 63)	.69 (.28)	.98 (.07)	.58 (.27)
	Above (n = 30)	.88 (.18)	.98 (.08)	.63 (.29)
Total		.76 (.27)	.98 (.08)	.60 (.27)
High (n = 180)	Below (n = 47)	.41 (.26)	.91 (.17)	.38 (.22)
	Average (n = 119)	.51 (.28)	.94 (.13)	.50 (.23)
	Above (n = 14)	.69 (.29)	.95 (.12)	.65 (.25)
Total		.50 (.28)	.93 (.14)	.48 (.24)

Discussion and General Conclusions

The major focus of this study was to characterize the grammaticality judgments of low-income AAE speakers whose language ability and dialect density vary along a continuum from low to high. The findings showed that dialect density and language ability both contributed significantly to children's grammaticality judgment performance with approximately 33% of the total variance accounted for by both predictors. Specifically, higher dialect use predicted lower standard scores and higher language ability predicted higher standard scores on the grammaticality judgment measure. The fact that dialect density influenced children's performance was not surprising considering the MC subtest includes items that are likely to be influenced by non-mainstream dialect. In particular, the measure includes items that are hallmark features of AAE such as lack of subject verb agreement with BE (e.g. The *was* all here for breakfast.) and multiple negation (e.g., We *don't* have *no* apples). As such, judging these types of items as correct most certainly contributed to the negative association between dialect use and overall performance on the MC subtest. Likewise, the fact that language ability predicted children's grammat-

icality judgment scores was also not surprising. Grammaticality judgments tasks index receptive grammatical knowledge which should undoubtedly be influenced by language ability and as indicated by regression analysis, as language ability increased, scores on the MC subtest increased as well. Taken together, these findings suggest that language ability and dialect density are separable constructs as they each account for unique variance and together they account for a third of the overall variance in grammaticality judgment performance.

Though the outcomes of the regression analysis specified a significant proportion of variance explained by dialect density and language ability, the results raise the question of what accounts for the other 67% of grammaticality judgment performance. Recall that the composite score encompassed five measures of phonological processing and one measure of vocabulary and syntax. Perhaps if the language composite was more representative of convergent measures of validity such as expressive morphosyntax, language ability would have accounted for even more of the variance in grammaticality judgment scores.

Another possibility is that working memory contributed to grammaticality judgment ability. Although grammatical judgment is a language skill, this task undoubtedly taps working memory in that children must hold a sentence in memory long enough to access their morphological knowledge and determine the sentences grammaticality. Noonan and colleagues (2014) reported that children with concomitant language and working memory impairments were less accurate in their grammatical judgments when the error occurred later versus earlier in the sentence indicating that the working memory demands imposed by later occurring errors compromised accuracy of grammaticality judgments. This error pattern was not evident for children who were classified as SLI only. In the current study children were classified as having below, average or above average language skills without regard for specific disabilities categories and it possible

that some children had deficits in working memory; however this data was not available for the participants in the current study. Future grammaticality judgments studies with this population should include measures of working memory as it is possible that this task presents additional cognitive demands for bi-dialectal speakers regardless of language ability.

When considering performance on all three item types, dialect groups differed in their responses to AAE items but not UNG items yet both items types distinguished performance among language ability groups. Similar to reports in the extant literature (Rice et al., 1999; Rice et al., 2009; Redmond & Rice, 2001), when considering the AAE items, children's grammaticality judgments paralleled their expressive language profile in that children who used AAE features in their speech infrequently were also less likely to judge AAE items as correct (or consistent with their grammar system) compared with their peers who used AAE more frequently and judged AAE items as correct more often. On the other hand, differences by language ability group, but not dialect group on UNG items confirmed that these items are indeed inconsistent with AAE. The UNG items consistently presented over-regularizations of past tense. Some researchers have suggested that over-regularized forms of past tense are characteristic of AAE (Horton-Ikard & Miller, 2004; Oetting & McDonald, 2001; Pruitt & Oetting, 2009). However, the results of the current study present evidence that contrasts with this claim. Of the six UNG items, four items included over-regularizations of past tense (e.g., The visitors com<u>ed</u> from the city.) and comparable performance between dialect groups indicates that these items function differently than the items classified as AAE. One possible explanation is that over-regularizations characterize the speech of young AAE speakers (Oetting & McDonald, 2001; Pruitt & Oetting, 2009), but this feature, as in MAE, may be developmental. For example, Horton-Ikard and Miller (2004) observed that seven- to 11-year-olds produced over-regularizations, but that the rate of

use decreased from age seven to nine, suggesting that this feature indeed may be developmental. "Over-generalizations" and "hypercorrections" coded by Horton-Ikard and Miller (2004) included both over-regularized forms of past tense and third person singular forms. The data presented here show that for AAE-speaking children eight- to 11-years-olds, over-regularized forms are not recognized as dialect forms, but may instead be recognized as developmentally appropriate for younger children who speak AAE making them ambiguous and difficult to judge for children the age of the participants in this investigation.

At first glance, it may seem surprising that AAE items were successful at distinguishing between language ability groups since some research on expressive morphosyntax has warned against and evaded the use of contrastive features in language assessment because they often overlap with clinical markers of impairment (Craig & Washington, 2000; Craig, Washington, Thompson-Porter, 1998; Seymour et al., 1998; Washington & Craig, 2004). This finding may partially be driven by the fact that all of the participants with below average language skills were also high dialect users and therefore, these children were more likely to be impacted by AAE-influenced items (either because of their dialect or low language skills or both). However, the results showed that children with average and above average language skills performed significantly different with the above average language group judging AAE items as incorrect more often; these groups included participants who were low and high dialect speakers. These findings contribute to the growing body of literature that supports the use of contrastive features in language assessment with this population (Garrity & Oetting, 2010; Oetting & McDonald, 2001; Oetting, McDonald, Seidel, & Hegarty, 2016), and suggests that the utility of contrastive features of AAE in the assessment process may be influenced by the density of dialect used.

Another goal of the current investigation was to examine a potential association between dialect density and overall language ability. As expected, the results support previous research that show children who use more AAE dialect features in their speech are more likely to demonstrate lower levels of performance on language measures compared with their peers whose speech more closely approximates MAE (Charity et al., 2004; Gatlin & Wanzek, 2016; Craig & Washington, 2004). It is important to keep in mind that language ability groups were classified relative to the current sample's distribution of scores, and while the data were normally distributed, the mean language composite was 8.08 (*SD* = 1.66) which is nearly two scaled points below the published normative mean of 10. Therefore, 1 *SD* of the mean as a cutoff for assigning language groups reflects that children who were classified as above average actually demonstrate average language skills when compared against the published normative data. In the existing language literature on low-income children, there is precedence for language scores to be normally distributed around a lower mean (Champion, Hyter, McCabe, 2003; Qi, Kaiser, Milan, Yzqueierdo, & Hancock, 2003; Restrepo et al., 2006; Washington & Craig, 1999).

Dialect density accounted for approximately 17% of the variance in language ability. It is likely that this negative, moderate correlation was driven, in part, by the below average language group, as 100% (47/47) of children in the below average group were high dialect speakers. Importantly, however, not all of the above average language users in the sample were low dialect speakers. In fact, 32% (14/44) of children with above average language ability were also classified as high dialect users, suggesting that high dialect use is not an a priori indicator of poor language outcomes. This finding is important and unexpected based on the current understanding of high AAE dialect use in low-income AA children. The preponderance of the evidence focused on dialect density has posited that low-income AAE-speaking children who are high dia-

lect users perform less well than their peer peers who are low or moderate users on a range of language-influenced tasks. There are a few studies that have found that young AAE speakers with high dialect density perform comparably to low dialect users and may exceed the performance of moderate dialect users. Connor and Craig (2006) reported that at risk, AA preschoolers with low and high dialect density outperformed their peers who used moderate amounts of dialect on measures of sentence imitation, phonological awareness, and letter-word identification. Similarly, in a study of first graders, Terry and colleagues (2010) also observed a U-shaped relationship with low and high dialect users outperforming moderate dialect users on measures of word reading and vocabulary; yet this relationship was moderated by SES. Children from higher SES schools outperformed their peers from lower SES schools. Thus, these findings suggest that dialect use is more likely to negatively impact language and literacy skills for children living in poverty. The current study emphasizes that there is variability in language ability even for children who are dialect users growing up in poverty. These differences cannot be attributed to varying exposure to MAE, as some researchers have concluded (Craig & Washington, 2004; Washington & Craig, 1998), because all participants in the current study lived in the same communities and attended low-income schools where FRPL eligibility was 87% or greater.

Future research should focus upon improving our understanding of the relationship between dialect density and language ability in this population. It may be especially informative to study the cognitive abilities in particular executive functioning of children who vary in their nonmainstream use. Research on bilingual, multilingual, and bi-dialectal children indicates a potential "cognitive advantage" for children who speak two or more languages, or dialects, compared to their monolingual peers (Antoniou et al., 2016; Barac, Bialystok, Castro, & Sanchez, 2014; Carlson and Meltzoff, 2008). This cognitive advantage has been evidenced in executive

function domains including inhibitory control, task-shifting, and working memory. It is not currently understood how executive function ability is associated with AAE use. Though it can be inferred from the extant literature that being a low dialect user is indicative of linguistic flexibility, which allows children to shift their language to match MAE (presumably the language used in classrooms and on standardized measures of achievement) affording them academic benefits over their high dialect-speaking mates. Data presented in the current study suggests that high dialect speakers may also be linguistically flexible and though it is beyond the scope of this paper, this is an area that deserves scholarly attention especially since executive functioning has been linked to language ability, behavioral outcomes, and academic achievement (Gutierrez-Clellan & Pena, 2001; Noble, Norman, & Farah, 2005; Qi & Kaiser, 2004; Schmitt, Justice, & O'Connell, 2014; Singer & Bashir, 1999). This line of research is clinically relevant and would contribute to the literature focused on improving language and literacy outcomes for AA children. Better understanding the relationship between dialect density and language ability is especially important as oral language skills are the foundation for reading and writing.

This study contributes to the extant literature base on AAE by providing a description of children's grammaticality judgments as a function of dialect density and language ability. Importantly, this study addresses a critical omission in the literature by exploring AA children's receptive language knowledge, indexed by their judgments of sentence level grammaticality. The implications of these findings suggest that grammaticality judgment tasks have the potential to be clinically useful for differentiating AA children with varying levels of language ability. In the current study, children's overall performance on the grammaticality judgment measures was significantly predicted by both dialect density and language ability. Furthermore, a moderate effect size was observed for language groups on AAE and UNG items suggesting

that these items may be informative for examining the divergence in linguistic knowledge between children with below and average language skills. These findings are limited by the fact that dialect density alone accounted for a significant portion of the variance explained in overall grammaticality judgment performance with differences by dialect group being notably evident on AAE items which suggests that the MC subtest of the TOLD-I:4 is not entirely sensitive to nonmainstream dialect use, in particular AAE. Participant responses to AAE items likely decreased overall standard scores: the high (M = 6.23) and low (M = 8.44) dialect groups earned scaled scores at the lower bounds of the average (8-12) and below average range (6-7), respectively. While selecting the MC subtest to measure grammatical knowledge allowed for participants to demonstrate knowledge of their dialect as well as MAE, this specific item composition limits the clinical utility of this tool for AAE speakers and, by implication, other nonmainstream dialect speakers. Nonstandardized assessments of grammatical judgment may be useful for future research.

The current study contributes to the growing body of literature on AA children that shows high dialect use is not automatically indicative of low language performance, as well as confirming the clinical utility of grammaticality judgments for use with AAE-speaking children. Importantly, this study provides unique, preliminary evidence with low-income, elementary-age children that indicates variability in language ability exists among dialect users at all levels. From these findings and others (Connor & Craig, 2006; Terry et al., 2010) researchers in the field should begin to change the narrative around nonmainstream dialect, moving away from a deficit perspective to addressing research with this population using a strength-based approach that focuses on the inherent abilities and skills these children bring to the table.

References

Antoniou, K., Grohmann, K. K., Kamanaros, M., & Katsos, N. (2016). The effect of childhood bilectalism and multilingualism on executive control. *Cognition, 149,* 18-30.

Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly, 29,* 699-714.

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11*(2), 282-298.

Champion, T. B., & Hyter, Y. D., McCabe, A., & Bland-Stewart, L. M. (2003). A matter of vocabulary: Performance of low-income African American head start children on the Peabody Picture Vocabulary Test-III. *Communication Disorders Quarterly, 24(1),* 121-127.

Charity, A. H., Scarborough, H. S., & Griffin, D. M. (2004). Familiarity with school English in African American children and its relation to early reading achievement. *Child Development, 75(5),* 1340-1356.

Connor, C. M., & Craig, H. K. (2006). African American preschoolers' language, emergent literacy skills, and use of African American English: A complex relation. *Journal of Speech, Language, and Hearing Research, 49,* 771–792. doi:10.1044/1092-4388(2006/055)

Craig, H. K., & Washington, J. A. (2000). An assessment battery for identifying language impairments in African American children. *Journal of Speech, Language, and Hearing Research, 43,* 366-379. doi:10.1044/jslhr.4302.36

Craig, H. K., & Washington, J. A. (2004). Grade-related changes in the production of African American English. *Journal of Speech, Language, and Hearing Research, 47,* 450–463. doi:10.1044/1092-4388(2004/036)

Craig, H. K., Thompson, C. A., Washington, J. A., & Potter, S. L. (2004). Performance of elementary-grade African American students on the Gray Oral Reading Tests. *Language, Speech, and Hearing Services in Schools, 35,* 141–154. doi:10.1044/0161-1461(2004/015)

Craig, H. K., Washington, J. A., & Thompson-Porter, C. (1998). Performance of young African American Children on Two Comprehension Tasks. *Journal of Speech, Language, Hearing and Research, 41,* 445-457.

de Villiers, J. G., & Johnson, V. E. (2007). The information in third-person /s/: acquisition across dialects of American English. *Journal of Child Language, 34,* 133-158. doi: 10.1017/S0305000906007768

Gatlin, B., & Wanzek, J. (2015). Relations among children's use of dialect and literacy skills: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 58,* 1306-1318. doi:10.1044/2015_JSLHR-L-14-0311

Garrity, A., & Oetting, J. (2010). Auxiliary BE production by African American English-speaking children with and without specific language impairmentf. *Journal of Speech, Language, and Hearing Research, 53,* 1307-1320. doi: 10.1044/1092-4388(2010/09-0016)

Gemake, J. S. (1981). Interference of certain dialect elements with reading comprehension for third graders. *Reading Improvement, 18*(2), 183-189.

Hammill, D. D., & Newcomer, P. L. (2008). *Test of Language Development-Intermediate: Fourth Edition.* Austin, TX: Pro-ed.

Horton-Ikard, R., & Miller, J. F. (2004). It is not just poor kids: the use of AAE forms by African-American school-aged children from middle SES communities. *Journal of Communication Disorders, 37,* 467-487. doi: 10.1016/j.jcomdis.2004.02.001

Johnson, V. (2005). Comprehension of third person singular in African American children. *Speech, Language, and Hearing in the Schools, 36*, 116-124. doi:10.1044/0161-1461(2005/011)

Johnson, K. R., & Simons, H. D. National Center for Educational Research and Development. (1973). *Black children's reading of dialect and standard text.* Retrieved from http://files.eric.ed.gov/fulltext/ED076978.pdf

Kaufman, A. S., Kaufman, N. L. (2004). *Kaufman Brief Intelligence test: Second Edition.* Bloomington, MN: Pearson.

Lee James, R., Washington, J. A., Seidenberg, M. (submitted). *Morphological judgments of low income, school-age AAE-speaking children.*

Miller, C., Leonard, L., & Finneran, D. (2008). Grammaticality judgments in adolescents with and without language impairment. *International Journal of Language and Communication Disorders, 43,* 346-360. doi: 10.1080/13682820701546813

Nelson, N. W., McRosky, R. L. (1978). Comprehension of Standard English at varied speaking rates by children whose major dialect is Black English. *Journal of Communication Disorders, 11*(1), 37-50. doi: 10.1016/0021-9924(78)90052-7

Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science, 8(1),* 74-87.

Noonan, N., Redmond, S., & Archibald, L. (2013). Contributions of children's linguistic and working memory proficiencies to their judgments of grammaticality. *Journal of Speech, Language, and Hearing Research, 57,* 979-989. doi:10.1044/2014_JSLHR-L-12-0225

Oetting, J., & McDonald, J. (2001). Nonmainstream dialect use and specific language impairment. *Journal of Speech, Language, and hearing Research, 44*, 207-223. doi:10.1044/1092-4388(2001/018)

Oetting, J., & McDonald, J. (2002). Methods for characterizing participants' nonmainstream dialect use in child language research. *Journal of Speech, Language, and Hearing Research, 45*, 505-518. doi:10.1044/1092-4388(2002/040)

Oetting, J. B., McDonald, J. L., Seidel, C. M., & Hegarty, M. (2016). Sentence recall by children with SLI across two nonmainstream dialects of English. *Journal of Speech, Language, Hearing Research, 59,* 183-194. doi:10.1044/2015_JSLHR-L-15-0036

Pawlowska, M., Robinson, S., & Seddoh, A. (2014). Detection of lexical and morphological anomalies by children with and without language impairment. *Journal of Speech, Language, and Hearing Research, 57,* 236-246.doi:10.1044/1092-4388(2013/12-0241)

Pruitt, S., & Oetting, J. (2009). Past tense marking by African American English-speaking children reared in poverty. *Journal of Speech, Language, and Hearing Research, 52,* 2-15. doi:10.1044/1092-4388(2008/07-0176)

Qi, C. H., & Kaiser, A. P. (2004). Problem behaviors of low-income children with language delays: An observation study. *Journal of Speech, Language, and Hearing Research, 47,* 595-609. doi:10.1044/1092-4388(2004/046)

Qi, C. H., Kaiser, A. P., Milan, S. E., Yzquierdo, Z., & Hancock, T. B. (2003). The performance of low-income, African American Children on the Preschool Language Sscale-3. *Ournal*

of Speech, Language, and Hearing Research, 46, 576-590. doi:10.1044/1092-4388(2003/046)

Redmond, S., & Rice, M. (2001). Detection of irregular verb violations by children with and without SLI. *Journal of Speech, Language, and Hearing Research, 44,* 655-669. doi:10.1044/1092-4388(2001/053)

Restrepo, M. A., Schwanenflugel, P. J., Blake, J., Neuharth-Pritchett, S., Cramer, E. S., & Ruston, H. P. (2006). Performance on the PPVT-III and the EVT: Applicability of the measures with African American and European American preschool children. *Language, Speech, and Hearing Services in Schools, 37,* 17-27. doi:10.1044/0161-1461(2006/003)

Rice, M., Hoffman, L., & Wexler, K. (2009). Judgments of omitted BE and DO in questions as extended finiteness clinical markers of specific language impairment (SLI) to 15 years: A study of growth asymptote. *Journal of Speech and Hearing Research*, *52*, 1417-1433. doi:10.1044/1092-4388(2009/08-0171)

Rice, M., Wexler, K., & Redmond, S. (1999). Grammaticality judgment of an extended optional infinitive grammar: Evidence from English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 43*, 943-961. doi:10.1044/jslhr.4204.943

Rickford, J. (1999). *African American Vernacular English.* Malden, MA: Blackwell Publishers Inc.

Schmitt, M. B., Justice, L. M., & O'Connell (2014). Vocabulary gain among children with language disorders: Contributions of children's behavior regulation and emotionally supportive environments. *American Journal of Speech-Language Pathology, 23,* 373-384. doi:10.1044/2014_AJSLP-12-0148

Seymour, H., Bland-Steward, L., & Green, L. (1998). Difference versus deficit in child African American English. *Language, Speech, and Hearing Services in the Schools, 29*, 113-138.

Seymour, H., Roeper, T., & de Villiers, J. (2003). *Diagnostic Evaluation of Language Variation—Screening Test*. San Antonio, TX: Harcourt Assessment.

Singer, B. D., & Bashir, A. S. (1999). What are executive functions and self-regulation and what do they have to do with language-learning disorders? *Language, Speech, and Hearing Services in Schools, 30,* 265-273. doi:10.1044/0161-1461.3003.265

Steffensen, M. S., Reynolds, R. E., McClure, E., & Guthrie, L. F. (1981, February). *Black English vernacular and reading comprehension: A cloze study of third, sixth, and ninth graders.* Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/17964/ctrstreadtechrepv01981i00199_opt.pdf?sequence=1

Terry, N. P. (2006). Relations between dialect variation, grammar, and early spelling skills. *Reading and Writing, 19,* 907-931. doi: 10.1007/s11145-006-9023-0

Terry, N. P. (2012). Dialect variation and phonological knowledge: Phonological representations and metalinguistic awareness among beginning readers who speak nonmainstream American English. *Applied Psycholinguistics,* 1-22. doi:10.1017/S0142716412000276

Terry, N., Connor, C., Thomas-Tate, S., & Love, M. (2010). Examining relationships among dialect variation, literacy skills, and school context in first grade. *Journal of Speech, Language, and Hearing Research, 53,* 126-145. doi:10.1044/1092-4388(2009/08-0058)

Terry, N. P., & Scarborough, H. (2011). The phonological hypothesis as a valuable framework for studying the relation of dialect variation to early reading skills. In S. Brady, D. Braze,

& C. A. Fowler (Eds.), Explaining individual differences in reading: Theory and evidence (pp. 97–117).

Tomblin, J. B., Records, N. L., Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 39,* 1284-1294. doi:10.1044/jshr.3906.1284

Thompson, C. A., Craig, H. K., & Washington, J. A. (2004). Variable production of African American English across oracy and literacy context. *Language, Speech, and Hearing Services in Schools, 35,* 269-282. doi:10.1044/0161-1461(2004/025)

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing-Second Edition.* Austin, TX: PRO-ED.

Washington, J., & Craig, H. (1994). Dialectal forms during discourse of poor, urban, African American Preschoolers. *Journal of Speech and Hearing Research, 37*, 816-823. doi:10.1044/jshr.3704.816

Washington, J. A., & Craig, H. K. (1998). Socioeconomic status and gender influences on children's dialectal patterns. *Journal of Speech, Language, and Hearing Research, 41,* 618-626. doi:10.1044/jslhr.4103.618

Washington, J. A., & Craig, H. K. (1999). Performance of at-risk, African American preschoolers on the Peabody Picture Vocabulary Test-III. *Language, Speech, and Hearing Services in Schools, 30,* 75-82. doi:10.1044/0161-1461.3001.7

Washington, J. A., & Craig, H. K. (2004). A language screening protocol for use with young African American children in urban settings. *American Journal of Speech-Language Pathology, 13,* 329-340. doi:10.1044/1058-0360(2004/033)

Washington, J. A., Craig, H. K., Kushmaul, A. J. (1998). Variable use of African American English across two language sampling contexts *Journal of Speech, Language, and Hearing Research, 41,* 1115-1124. doi:10.1044/jslhr.4105.1115

Wulfeck, B. (1993). A reaction time study of grammaticality judgments in children. *Journal of Speech and Hearing Research, 36,* 1208-1215. doi:10.1044/jshr.3606.1208

Wulfeck, B., Bates, E., Kwiatkowski-Krupa, M., & Saltzman, D. (2004). Grammaticality sensitivity in children with early focal brain injury and children with specific language impairment. *Brain and Language, 88,* 215-228. doi:10.1016/S0093-934X(03)00100-7

APPENDIX A

Items Selected from the MC subtest of the TOLD-I:4

1. Those boys is happy. AAE
2. We maked some pudding. NMAE
3. We haven't no candy to give her. NMAE
4. They like school. MAE
5. He done the dishes already. AAE
6. They was all here for breakfast. AAE
7. He drinked the cola. NMAE
8. Carla don't like school. AAE
9. It don't matter what you say. AAE
10. He is tall. MAE
11. She's the faster runner of them all. NMAE
12. Marcus don't run fast. AAE
13. She growed three inches this year. NMAE
14. The visitors comed from the city. NMAE
15. We don't have no apples. AAE
16. Them two boys joined the navy. AAE
17. Betty watches TV. MAE
18. Them girls stole the cake. AAE