The source of Black–White inequality in early language acquisition: Evidence from Early Head Start

Daniel Kreisman

Social Science Research

Social Science Research 41 (2012) 1429–1450

Contents lists available at SciVerse ScienceDirect

Social Science Research

journal homepage: www.elsevier.com/locate/ssresearch

The source of Black–White inequality in early language acquisition: Evidence from Early Head Start

Daniel Kreisman

Harris School of Public Policy, University of Chicago, Chicago, IL 60637, United States

ARTICLE INFO

Article history:
Received 22 October 2011
Revised 23 February 2012
Accepted 10 May 2012
Available online 31 May 2012

Keywords:
Language
Race
Home language environment
Factor analysis
Decomposition
Head Start
Growth model

ABSTRACT

I compare language learning trajectories for Black and White children over the first 3 years of life using data from the Early Head Start Research and Evaluation study (EHSRE) in order to determine the timing and source of divergence in early language skill. Results indicate that that while controlling for racial differences in family background and a measure of the home language environment cannot entirely account for disparities in language acquisition, interactions between age, race and maternal education, and between race and a time-varying measure of the home language environment play a significant role. I show that returns to parental education and the home language environment, in terms of language learning, are higher for White children than their Black peers. Specification checks confirm that these results are robust to alternate definitions of child language and the home language environment, and that no interactions between race, age and any of the other covariates are significant. I discuss possible explanations for these race specific education gradients, including measurement error and test bias. In addition, I address relevant empirical issues in estimating language growth with respect to linguistic inputs and the home language environment.

1. Introduction

That Black children on average trail their White peers in reading and language achievement by school entry is well documented (Brooks-Gunn et al., 1996; Fryer and Levitt, 2004; Magnuson and Duncan, 2006), yet the source and timing of divergence in linguistic ability is not as well understood. Although several studies address the relationship between SES and early language (Hart and Risley, 1995; Hoff, 2003; Rowe, 2008; Weizman and Snow, 2001), few have done so with diverse samples and fewer yet focus on the source of differences by race. Most language studies are concerned with the relative timing of various inputs in the development process (Hoff and Naigles, 2002; Huttenlocher et al., 1991, 2007; Pan et al., 2004, 2005). Intuition suggests that sufficiently accounting for these inputs, initial conditions, family circumstances and other observed variables that affect language learning should largely explain differences across race. Failure to do so indicates the existence of unobserved heterogeneity affecting early skill development.

Rodriguez et al. (2009) and Rodriguez and Tamis-LeMonda (2011), using the same data as the present study, show that Black mothers fare worse on measures of the home language environment and that Black children perform worse on tests of language skill between 14 and 36 months, but do not identify the source of this disparity. Similarly, Farkas and Beron (2004) find a Black–White gap in language at 36 months, their baseline observation, and that controlling for SES and the home environment cannot fully account for this. They claim that returns to maternal inputs and maternal education, in terms of child language learning, accrue for White children before 36 months, but not until after 36 months for Black children. Yet, this result is inferred from birth to age 36 months as the authors' data begins at age 3. In the following I explicitly test this hypothesis

E-mail address: dannyk@uchicago.edu

0049-089X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ssresearch.2012.05.010

using data from the Early Head Start Research and Evaluation study (EHSRE), which contains measures of child and parental language at 14, 24, and 36 months. I interact a time varying measure of the home language environment with race, and interact various measures of socioeconomic status (SES) with race and age to test for differential rates of growth. I then ask whether the emergence of the Black–White language gap is due to differences in family background characteristics and the home language environment, or if the relationship between these characteristics and early language acquisition differs across race by estimating a standard decomposition.

Results indicate several significant findings. First, a substantial Black–White gap in in the home language environment, measured by a principal factor score from linguistic variables from the HOME and 3-bag evaluations of the child's learning environment, is apparent at 14 months and remains relatively constant through age 3. Second, although at 14 months Black children slightly outperform their White peers in language, again measured by a principal factor score, by 24 months a significant gap in language emerges which widens by 36 months; these gaps persist despite a robust set of controls for demographic characteristics, family circumstances and the home language environment. Third, persistent negative coefficients on interactions between age, race and maternal education and race and a time-varying measure of the home language environment confirm Farkas and Beron's (2004) hypothesis that the returns to measures of the home language environment accrue differently with respect to race over these ages. Lastly, a decomposition of the gap at each wave indicates that while at 24 months differences in endowments of the covariates explain much of the measured language gap, at 36 months only half of the race gap is due to differences in endowments, and half is attributable to differences in returns to these endowments, in particular to differential returns to maternal education and a measure of the home language environment. Robustness checks, including alternate measures of the home language environment, interacting age, race and all covariates, and re-estimating the main specification using all possible combinations of child language measures at each age as opposed a single factor score, confirm these findings. In addition, I discuss and test for bias from simultaneously determined parental language inputs and child language outcomes.

This study presents two primary limitations. First, a significant share of respondents do not have valid home and child language measures at each age; robustness checks verify that this does not substantively alter results. Second, in order to observe a large and diverse sample, child language and the home language environment are measured using prescriptive measures, in several cases maternal reports of child language, rather than an actual record of the quantity and type of child and maternal speech. Both limitations and benefits from this method are discussed in Section 2.

The remainder of the paper is as follows: Section 2 presents relevant theory and a summary of the literature, Section 3 describes the data and estimation sample, Section 4 describes the home language environment and child language factor score measures, Section 5 describes the estimation strategy, and discusses results and Section 6 summarizes with suggestions for further research.

2. Background

2.1. Child language, maternal speech and SES

Beginning with Wolfram (1971) and Labov (1972), a rich literature evaluates differences in first language acquisition patterns across race and social class.[1] Such studies often focus on estimating the effect of variation in maternal speech on child language development by recording and analyzing dyadic interactions between mothers and their children in various settings. The implication is that the primary mechanism through which race and SES affect language is the home language environment.

Several early studies demonstrate clear differences in maternal speech patterns across socioeconomic strata. Analyzing interactions from 30 working class and 33 upper class mother–child dyads in various settings, Hoff-Ginsberg (1991) finds significant differences across social class in mothers' child directed speech and moderate differences in mothers' adult directed speech, suggesting that general linguistic differences in speech across socioeconomic strata parallel differences in language development among children. Hart and Risley (1995) identify considerable disparity across class in the amount of words children hear by analyzing monthly observations of parent–child dyads in 42 families. The authors show that divergence in language acquisition across class begins as early as 24 months and that by 36 months large gaps emerge.

While early studies focus on quantity and proportions in maternal speech, subsequent studies shift toward measures of maternal speech quality demonstrating that child-directed maternal speech changes as children develop. Weizman and Snow (2001), using a corpus of nearly 150,000 speech tokens from 53 low income mothers, find that nearly 99% of maternal inputs in the study were comprised of the 3000 most frequent words. They conclude that the density of the remaining "sophisticated" words children heard predicted over a third of the variation in child vocabulary in second grade. Pan et al. (2005), using a sample of 108 low and middle class families, claim that variation in quantity of maternal speech is not the strongest predictor of variation in child vocabulary acquisition and that other factors, such as nonverbal communication, linguistic diversity or social context may play equal or greater roles in determining language development. Rowe et al. (2005), using a sample of 108 mother–child dyads, show that maternal speech increases in both amount and diversity as children age, and that although maternal characteristics such as education, language skill and depression explained variation in maternal speech, they were unable to entirely explain variation in child language growth.

[1] More recent studies include: Clopper and Pisoni (2004), Bailey and Thomas (1998), Rickford and Rafal (1996), Washington and Craig (2002), and Wolfram (1991).

2.2. Education gradient

These findings coincide with a broader literature on parental time use and the economics of the family (Becker, 1965) in which parents allocate time between household production, such as childrearing, work, and if possible, leisure. Working within this framework, Guryan et al. (2008) posit that more educated parents will spend more time with their children, finding evidence for this "education gradient" in the American Time Use Survey (ATUS). Kalil et al. (forthcoming) extend this theory, suggesting that more educated mothers will not only spend more time with their children on average, but will also more suitably tailor their interactions to the appropriate developmental stage. Using the ATUS, Kalil et al. show that in addition to spending more time with children overall, more educated mothers spent significantly more quality time with their children. In particular, between ages 0 and 2 college educated mothers spent more time in basic child care, at play and in teaching activities than their less educated peers, with little difference in time spent on household "maintenance". Between ages 3 and 5, college educated mothers increased the share of time they spent on teaching activities, contributing to a widening disparity between college and non-college educated mothers in instructional time spent with children. The authors conclude that the widening skill gap is in part due to these "education gradients" in both the amount and type of time parents spend with children. While neither Kalil et al., nor Guryan et al., test for these gradients by race, Kalil et al. find that the main effect for race (both Black and Hispanic) is negative.

2.3. Race and language development

Similar to findings here, Rodriguez et al. (2009) show that the language gap is nonexistent or favoring Black children at 14 months, and that it widens, favoring White children, by 36 months. The authors do not assess why growth rates are different for Black and White children after controlling for background characteristics or the home environment. Rodriguez and Tamis-LeMonda (2011), again using data from the EHSRE, classify families by growth in the home language environment between ages 1 and 5 to see how this explains child language growth between ages 3 and 5. They show that children in homes with positive home learning environment "trajectories" performed better on a language test at pre-k and kindergarten, controlling for background characteristics. They also show that Black children perform worse on the outcome test and that Black families are more likely to have low or declining home language environment trajectories than their White peers. They do not assess whether controlling for these characteristics explains the race gap or whether effects accrue differently for Black or White families. They do show that measures of the home language environment provide a significant amount of explanatory power above and beyond basic demographic controls in predicting child language development.

Farkas and Beron (2004) address race and language development using a larger sample in the Children of the National Longitudinal Survey of Youth (CNLSY79) to estimate differences in vocabulary growth rates across race and SES during preschool and the first years of schooling. The authors find a gap between Black and White children in language ability at 36 months, their earliest observation, and note that very little of the early gap can be explained by SES. They find that the relative size of the race gap remains unchanged through age 13 and that the effect of SES on vocabulary growth for White children occurs almost entirely before 36 months while for Black children the effect of SES is concentrated between 3 and 4 years. Importantly, the authors note that gaps in language ability between Black and White children, controlling for SES, emerge by 36 months and persist through age 13. The authors control for maternal inputs using the HOME evaluation and a measure of mother's verbal ability from the Armed Forces Qualification Test (AFQT) administered to NLSY participants. They conclude that social class explains little of the Black–White gap in language growth before 36 months and that mother's verbal AFQT score and the HOME score can account for only about 35% of this gap.

Pungello et al. (2009), using a sample of 146 children, estimate the effects of mother–child linguistic interactions observed at 12 and 24 months on child language between 18 and 36 months for Black and White children. The authors find that at 18 months measured expressive communication is similar for Black and White children, and that a gap emerges by the 24th month. Although the authors attempt to separate the effect of race from SES (in part by excluding the highest and lowest SES families from analysis), their sample is relatively small and they note that conclusions concerning SES and race should be interpreted with caution.

2.4. Empirical approaches to estimating differences in language development: metric and dyadic approaches

The studies mentioned above chronicle two distinct approaches to estimating language development in recent literature. The first approach uses observations of linguistic interactions between mother–child dyads at various ages to estimate the relationship between maternal inputs and child language development. This *dyadic approach* incorporates immediate measures of parental and child speech, such as the number of word types or tokens spoken by parent and child, or similar measures such as the mean length of utterance, number and type of gesture, or the complexity of maternal speech. Although these studies elicit the subtleties of the relationship between maternal speech and child language, the level of detail they require restricts sample size, prohibiting accurate inference across groups such as race or SES.

A second, less common approach is to administer normed tests either directly to the child, or using the mother as a proxy respondent, to measure the child's language development. Common measures include the MacArthur Communicative Development Index (Fenson et al., 2000) and the Peabody Picture Vocabulary Test amongst others (Dunn and Dunn, 1997). Although this *metric approach* does not provide the level of detail that the *dyadic approach* affords, it maintains two clear advantages. First, it allows researchers to use significantly larger and more diverse samples, especially with respect to income, race and SES, with richer sets of control variables. Second, measuring the language of mothers and their children in a dyadic setting will bias estimates of the effect of language inputs on child language outcomes, as well as other covariates, if children and mothers elicit speech from one another during measurement (this is discussed in detail in Section 4). Using an instrument to evaluate speech that abstracts language measurement from the immediate parent–child interaction eliminates this source of bias and allows researchers to treat language, in this specific sense, as exogenous.

Although the comparability between the two methods has not been explicitly tested, Pan et al. (2004), using a subsample from the data presented here, compare various test-based measures of child language, such as the MacArthur CDI, elements of the Bayley Scales of Infant Development, and the Peabody Picture Vocabulary Test, with detailed measures such as the number of types and tokens produced by the child. The authors find correlations ranging from 0.49 to 0.66 between measures of child types and tokens at 24 months and the MacArthur CDI and items from the Bayley Scales respectively. Correlations between the MacArthur CDI and Bayley Language items at 24 months, and between MacArthur and PPVT at 36 months, were higher than correlations between the PPVT and measures of child types and tokens at 24 months.[2]

Ultimately, the appropriate measure of language inputs depends on the goal of analysis. Where tokens measures a child's productivity, they fail to measure diversity, a problem remedied by measuring types. But, language is not only productive and neither of these measures directly account for receptive language. Hoff-Ginsberg (1992) discusses the significance of using both the total amount of speech children hear and the proportional frequency of various forms of maternal speech when evaluating the relationship between maternal speech and child language learning. The author highlights inherent difficulties in measuring what she describes as, "the child's actual experience."[3] Thus, despite the drawbacks of using tests to measure child language rather than counts of tokens or types – that they are less precise measures of language – they can provide a more clear picture of the child's overall language ability.

3. Data

3.1. The Early Head Start Research and Evaluation study

The national Early Head Start Research and Evaluation study (EHSRE), initiated in 1995, was designed to evaluate the impacts of Early Head Start programs on low income families (Love et al., 2002, 2005). Three main program types were evaluated: home-based, center-based, and a mix of home and center-based Head Start services in 17 locations throughout the US. Although the present study makes use of restricted data, which allows for item-level analysis of child language and home environment measures, these files do not include specific site or region identifiers making an evaluation of the program's effectiveness tenuous. Recently released pre-k and kindergarten waves were not available at the time of writing.

The sample is composed of low-income families with children up to 12 months in age at the time of enrollment; up to 10% of space was eligible for families with incomes above the poverty line. The initial sample included 3001 families, 1513 of whom were randomly assigned to the treatment group with 1488 assigned to the control group, meaning they could not receive Early Head Start services until after the child's 3rd birthday, but could enroll in other child care services. Baseline data was collected at enrollment and includes most demographic information. Parent service follow-up interviews, which collected data about child care services along with information about parental employment, social services received and family and child health, were conducted at 6, 15, and 26 months. At 14, 24, and 36 months, parent interviews and child and family assessments were conducted to evaluate the focus child's development and to track various family and developmental indicators; these comprise the main source of data presented here.

3.2. Analysis sample

From the original sample of 3000, 2799 observations are in the final dataset. Of these only Black and White children from English speaking households with a complete set of applicable sociodemographic data and with child language, HOME score and 3-bag evaluations in at least one wave are retained. Table 1 shows how I arrive at the working sample of 1458. Of these, 235 observations are missing poverty ratio data; the poverty ratio for these observations is set to zero and a dummy variable indicates missing data. As a check, I include measures indicating WIC and welfare recipient, neither of which alter the results thus they are not included in the final analysis. Keeping only observations with complete data in all waves results in a sample of 426, which reduces representativeness and increases the likelihood of bias from sample attrition.

Table 2 shows demographic characteristics by race. White families are wealthier and White mothers are more likely to have completed schooling past high school, and are less likely to have dropped out. Although Black and White respondents are not identical in all categories, there is considerable common support resulting from low-income eligibility requirements for Head Start participation. Table A1 in the Appendix breaks down sample demographics by data availability in each cross-section, those used in regressions with lagged measures as independent variables, and those in all waves. The three cross-sectional samples, at 14, 24, and 36 months, include observations from the working sample of 1458 with all maternal speech

[2] pp. 597–598.
[3] p. 242.

Table 1
Sample restrictions.

Sample restriction	*N*
Full sample	2977
Black and White	1983
English spoken in home	1951
Non-missing covariates[a]	1933
Child and home language in any wave	1458
Data in all three waves	426

[a] 32 are missing mother's HGC, 13 are missing gender, 3 are missing age at 1st birth, 2 are missing birth order.

Table 2
Summary statistics by race.

	White		Black		Total	
	Mean	*sd*	Mean	*sd*	Mean	*sd*
Boy	0.49	(0.50)	0.52	(0.50)	0.51	(0.50)
Poverty ratio	0.64	(0.46)	0.52	(0.54)	0.59	(0.50)
Poverty missing	0.10	(0.31)	0.23	(0.42)	0.16	(0.37)
HGC < 12	0.32	(0.47)	0.47	(0.50)	0.39	(0.49)
HGC > 12	0.30	(0.46)	0.24	(0.43)	0.27	(0.45)
Light birth	0.05	(0.22)	0.08	(0.26)	0.06	(0.24)
First born	0.60	(0.49)	0.70	(0.46)	0.65	(0.48)
Siblings age 0–5	1.19	(0.78)	1.11	(0.82)	1.15	(0.80)
Age first birth	22.73	(5.43)	21.01	(5.35)	21.92	(5.46)
Live with husband	0.36	(0.48)	0.07	(0.26)	0.22	(0.42)
Depression	2.34	(1.24)	2.94	(1.11)	2.62	(1.21)
Home based	0.50	(0.50)	0.32	(0.47)	0.41	(0.49)
Center based	0.19	(0.39)	0.28	(0.45)	0.23	(0.42)
Treatment	0.52	(0.50)	0.51	(0.50)	0.52	(0.50)
N	774		684		1458	

Sample from EHSRE includes all Black and White native English speakers with complete HOME and 3-bag evaluations in at least one wave. Omitted categories are HGC = 12 and Mixed Head Start treatment.

measures and child language measures in that particular wave. The 965 observation *lagged* sample includes observations with complete information at either 24 or 36 months with a lagged measure of the home language environment; Table A2 in the Appendix shows summary statistics by race and mother's HGC.

4. Measures

Ideally researchers would have detailed measures of the type and quantity of child directed maternal speech along with various detailed measures of a child's language ability throughout childhood. This takes a great deal of resources and remains, for the moment, unfeasible for a sample large enough make claims about group differences in the greater population. Thus, many large scale data sets contain composite measures of parental inputs and child outputs. In the following, for the home language environment, I make use of two measures in the data: the HOME evaluation and the 3-bag assessment. For child language outcomes, I use elements from the MacArthur Communicative Development Inventory at 14 and 24 months, items from the Bayley Scales of Infant Development at 24 months, and the PPVT at 36 months.

Various strategies are employed throughout the literature to make use of multiple measures of a small number of latent variables. Rodriguez et al. (2009) create classifications (low, moderate and high) for three separate domains of parenting quality: literacy activities, quality of maternal engagement, and provision of learning materials. For example, parents who read to children rarely and who told stories rarely are classified as low. Reading a few times a week or telling stories a few times a week is moderate, and conducting either of these activities daily places respondents in the high category. Measures, and therefore classifications, vary across waves. A similar strategy is employed by Rodriguez and Tamis-LeMonda (2011). In this case the authors create a score of 0–2 for each of the domains listed above and use the composite sum (0–6) at each age. Lugo-Gil and Tamis-LeMonda (2008) create a composite score for a single measure of parenting quality by summing up *Z*-scores created from measures of parenting activities such as the HOME score and 3-bag task to create a single measure of the home language environment. In each case the authors use confirmatory factor analysis to ensure that the measures load onto a single factor. Pungello et al. (2009) average scores from observations of parenting quality at 14 and 24 months using a factor analysis to identify relevant items. Farkas and Beron (2004) use unscaled composite measures from the HOME evaluation directly in their regression.

I take more flexible approach to measuring the home language environment at each wave. Rather than using composite scores for home environment measures, which place equal weight on each item, I estimate factor scores from items that are comparable over the three waves and which measure some dimension of the language children are exposed to in the home.

Table 3
Factor loadings, by wave.

	14 months		24 months		36 months	
	Loading	Uniqueness	Loading	Uniqueness	Loading	Uniqueness
3-Bag items						
Parental supportiveness	0.64	0.59	0.62	0.62	0.60	0.64
Parental detachment	0.49	0.76	0.48	0.77	0.37	0.86
Parental intrusiveness	0.35	0.87	0.36	0.87	0.39	0.85
Parental negative regard	0.38	0.86	0.40	0.84	0.45	0.80
HOME items						
10 books present and visible	0.26	0.93	0.22	0.95	0.30	0.91
Times read stories to child in past month	0.18	0.97	0.19	0.96	0.20	0.96
How many children's books	0.23	0.95	0.14	0.98	0.21	0.96
Parent spontaneously praises child	0.63	0.61	0.57	0.68	0.55	0.70
Parent voice conveys positive feeling	0.45	0.80	0.43	0.81	0.57	0.67
Parent names objects	0.53	0.72	0.44	0.81		
Spontaneously vocalized to child	0.59	0.65	0.41	0.83		
Parent responds to child speech	0.57	0.67	0.51	0.74		
Mother converses with child at least twice					0.64	0.59
Mother answers child's questions verbally					0.54	0.70
Mother uses complex sentence structure					0.43	0.81
Parental speech distinct	0.42	0.83	0.43	0.81		
Parent initiates verbal exchange with visitor	0.63	0.60	0.61	0.63		
Parent converses freely and easily	0.60	0.64	0.64	0.59		

Results from unrotated principal factor analysis using all Black and White English speaking families with completed HOME and 3-bag evaluations in a given wave. Factor loadings show the correlation between each item and the principal factor. Uniqueness indicates the variance in each item that is not accounted for by the other variables.

The advantage to this approach is that it imposes less structure on the weight each item has in the composite measure and allows the weight of each item on the factor score to vary across waves.

4.1. Linguistic input measures

4.1.1. The Home Observation for Measurement of the Environment (HOME)

At each wave researchers collected information from the Home Observation for Measurement of the Environment Bradley, which measures "the quality of stimulation and support available to a child in the home environment" (Love et al., 2002). Items for the HOME evaluation are scored by interviewers through a combination of parent interviews, and from the interviewer's own observations of the home environment and interactions between parent and child. At 14 and 24 months, the HOME evaluation includes the following subsections: Emotional Responsivity (Parental Warmth), Language and Cognitive Stimulation, and Parental Verbal/Social Skills. At 36 months the evaluation includes a Parental Warmth scale and Language and Cognitive Stimulation, but does not include Parental Verbal/Social Skills. In the public-release data, missing sub-scale items are imputed to the sample mean when less than 25% of items are missing for a given observation. This induces measurement error and will bias coefficients if missing observations are not random. To remedy this factor scores are estimated using only observations with all relevant items at each wave. Several items from the Language and Cognitive Stimulation subsection, particularly those relating to the child's number and type of toys, were excluded from analysis as these instruments are also proxies for family wealth and are not directly related to maternal speech inputs.[4]

4.1.2. The 3-bag assessment

In a separate evaluation, the parent and child were given three bags of toys to play with. The interactions were videotaped and coded by child development researchers. The ratings measure various aspects of parenting during semi-structured play. All items are on a seven point scale and are recoded such that 1 indicates a low incidence of a certain behavior and 7 indicates a high incidence. Parental supportiveness provides a measure of parental sensitivity, cognitive stimulation and positive regard toward the child during play. Parental detachment measures the extent to which the parent is inattentive, perfunctory or cold when interacting with the child. Parent intrusiveness measures the extent to which the parent fails to recognize and respect the validity of the child's perspective. Lastly, parent negative regard evaluates the degree to which the parent uses a negative tone or shows frustration or anger toward the child (Love et al., 2002; Ware et al., 1998).

4.1.3. Home language factor scores

Table 3 shows the relevant factor loadings and available items by wave; Fig. A1 in the Appendix plots the distribution of the Home language factor score by race and wave. The factors are estimated by including all items in Table 3 and estimating a principal component factor using the entire available sample with these measures. The factor loadings show the correlation between each item and the principal factor. The uniqueness indicates the variance in each item that is not accounted for by

[4] See Love et al. (2002) and Ware et al. (1998) for details.

the other variables. Variables with greater uniqueness have, by construct, less weight in the factor score. The items that most consistently load heavily on the home language factor are those that indicate the type, quantity, quality or milieu of the mother's speech.

4.2. Child language measures

Several measures of child language exist at 14 and 24 months, although only the PPVT exists at 36 months. Factor loadings for *child language* are constructed in the same manner as the *home language* factor score at each wave. Fig. A2 in the Appendix plots the distribution of child language factor scores by race and wave; the following measures are used to construct this variable.

4.2.1. MacArthur Communicative Development Inventory (CDI)

At 14 months the MacArthur CDI includes 89 word items. Parents are first asked if their child understands the word; if the child understands the word the respondent is then asked if the child also says the word. The raw score is computed separately for comprehension and production, each ranging from 0 to 89 words. Parents are then asked if their child performs any of a series of 12 gestures often, sometimes, or not at all. Gestures include items such as pointing and reaching at objects, waiving, nodding yes or no, or indicating "all gone." At 24 months, parents are again asked about their child's vocabulary production from a list of 100 words, with no questions about comprehension. Parents are then asked a series of 36 questions about the child's ability to use grammar or make sentences. For each item, the interviewer reads two phrases or word combinations, one simple and one more complex, and asks the parent which most resembles how her child speaks. These items illicit a measure of the child's complexity of speech and grammar usage at 24 months. Children who were not yet able to combine words are given a score of zero.

4.2.2. Bayley Language Factor

At 24 months, a subsection of the Bayley Scales of Infant Development (Bayley, 1993) measures the child's language ability. This Bayley Language score is the sum of 12 language items from the Bayley test that resulted from factor analysis conducted by the principal investigators; the specific items are listed in the Appendix (see Love et al. (2001) for further information).

4.2.3. Peabody Picture Vocabulary Test (PPVT)

At 36 months respondents were administered the PPVT, which measures receptive vocabulary and listening comprehension of spoken words. Children were given several series of pictures and asked to point to the picture that matched the words spoken by the interviewer. The test is age normed to a nationally representative sample.

4.3. Child and home language gaps over time

Fig. 1 plots the mean White–Black gap for the resulting child language factor score and the home language factor score across waves. Included are White–Black gaps in each of the individual child language test scores. Figs. 2 and 3 show Black–White gaps in child language and the home language factor respectively (normalized over the whole sample), by mother's highest grade completed. In Fig. 2 it is clear that for White children, while little difference in growth over time exists between children of mothers with less than high school and those with a high school degree, children of mothers with schooling beyond high school see rapid growth over the relevant time frame. For Black children, a large gap emerges for respondents of mothers with less than a high school degree. Fig. 3 plots home language scores by age across maternal education. This measure of the home language environment increases with maternal education for both Black and White children, yet home language scores for Black children are significantly lower than those of White children across maternal education. Over the 22 month period there is relatively little change in slope for any group. Table A3 in the Appendix shows pairwise correlations between all language measures and the home language and child language factor scores.

4.4. The relationship between inputs and language measures

Regressing measures of child language ability on a set of controls that includes contemporaneous maternal inputs will bias estimates if the mother's speech is in any way responsive to the child's language. This might result in two cases: if child language and maternal speech are simultaneously determined, insofar as mothers adjust speech to match that of the child, or if children and mothers elicit speech from one another during measurement. The latter case is not an issue in these data as child language measures are constructed from normed tests while maternal language is measured by trained observers at a separate time. In the former case, maternal language increases in response to the child's own language growth. Consider a cross-sectional estimate of maternal inputs and family characteristics on child language:

$$Y_i = \alpha_0 + \alpha_1 X_i + \alpha_2 Z_i + \alpha_3 Age_i + \epsilon_i, \quad (1)$$

where y_i is a measure of child language, X_i is any fixed characteristic such as race or SES, Z_i is some measure of child directed maternal speech, and ϵ_i is an error term containing all unobserved information. We expect that the coefficient on maternal

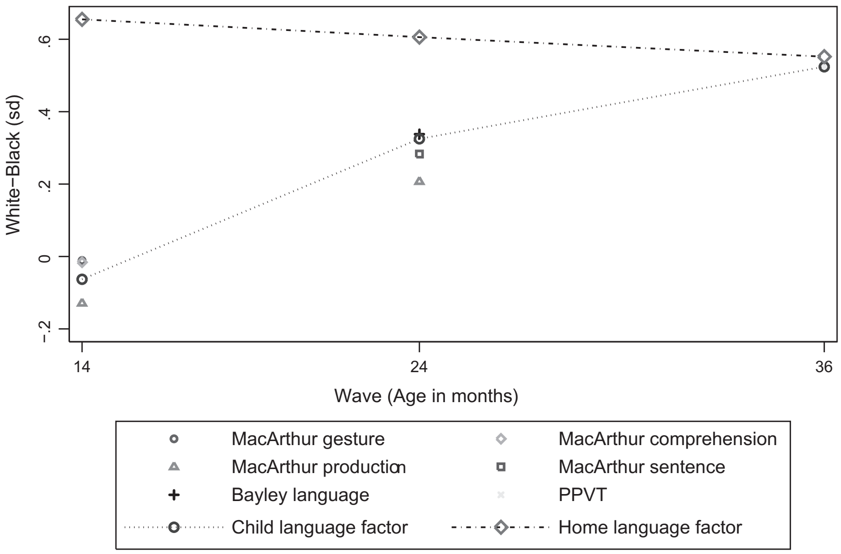

Fig. 1. White–Black gaps in child language and the home language environment at each age. All measures are normalized to a *Z*-score using the entire available sample. *Y*-axis is $\text{Mean}_{Black}-\text{Mean}_{White}$ for each measure. MacArthur Gesture and Comprehension are at 14 months only, Bayley language is at 24 months only and PPVT is the only score at 36 months. MacArthur production is measured at 14 and 24 months.

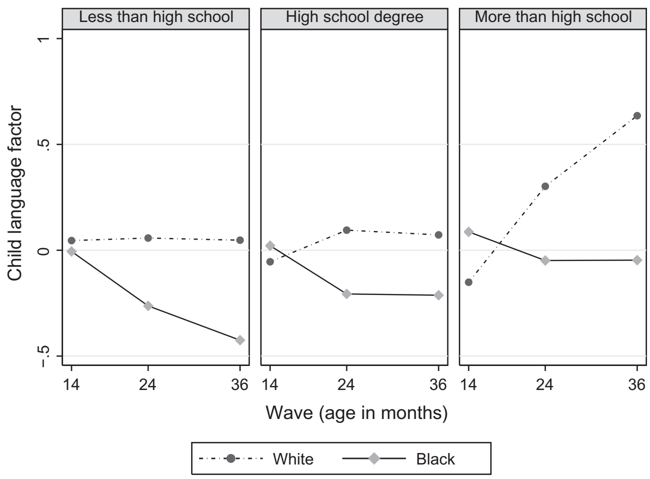

Fig. 2. White–Black gaps in child language factor scores, by race and maternal education. Plots are mean child language factor scores, normalized to *Z*-scores, by wave (age in months). Sample includes all cross-sectional observations at each wave with both child language and home language factor scores.

inputs, α_2, is positive. But, we might also believe that how mothers speak to their children is influenced by the child's speech as well. Rowe et al. (2005) and Rowe (2008) demonstrate that child word types have explanatory power in predicting child directed maternal speech. Thus, one could also estimate the effect of child speech on maternal speech:

$$Z_i = \beta_0 + \beta_1 X_i + \beta_2 Y_i + \beta_3 Age_i + \mu_i. \tag{2}$$

Plugging Eq. (1) into Eq. (2) and simplifying yields:

$$Y_i = \frac{(\alpha_0 + \alpha_2\beta_0)}{(1-\alpha_2\beta_2)} + \frac{(\alpha_1 + \alpha_2\beta_1)}{(1-\alpha_2\beta_2)}X + \frac{(\alpha_3 + \alpha_2\beta_3)}{(1-\alpha_2\beta_2)}Age_i + \frac{(\epsilon + \alpha_2\mu)}{(1-\alpha_2\beta_2)} \tag{3}$$

$$= \pi_0 + \pi_1 X_i + \pi_2 Age_i + \eta_i. \tag{4}$$

Thus, estimating the effects of *X* on child language in Eq. (1), for example race or SES, will result in a biased estimate of α_1 in π_1. Similarly, the growth rate of child language, α_3, or any interactions between the *X*'s and Age used to estimate differential growth rates will be biased as well. This is a standard result from estimating equations where inputs and outputs are simultaneously determined. Using the same data as the present study, Lugo-Gil and Tamis-LeMonda (2008) address the former case using a structural model with respect to family resources, parenting quality and general cognitive skill. They test

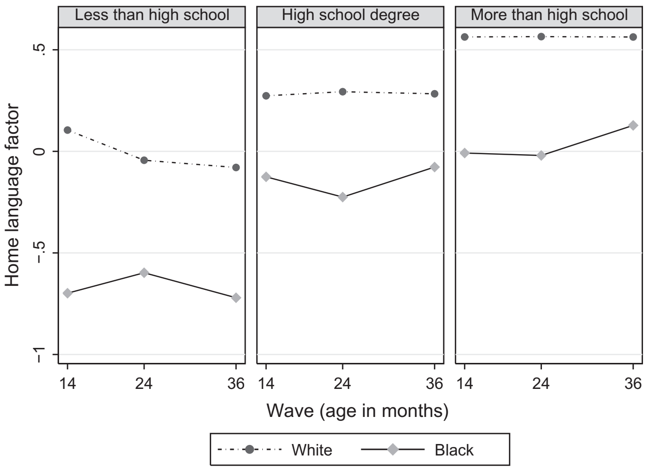

Fig. 3. White–Black gaps in home language factor scores, by race and maternal education. Plots are mean home language factor scores, normalized to Z-scores, by wave (age in months). Sample includes all cross-sectional observations at each wave with both child language and home language factor scores.

whether lagged versions of parenting quality are associated with current measures of child skill and whether past measures child skill are associated with current measures of parenting quality. They find that reciprocal effects exist and include all pathways in their model.

The second case is more specific to attempts at estimating a causal relationship between maternal speech and child language in which child and maternal speech are measured during a dyadic interaction. Although threat of bias from the previously described scenario still applies, this second, independent estimation problem exacerbates the measurement issue by overstating the relationship between linguistic inputs and outputs. I avoid the latter issue as child language here is not measured during dyadic interactions. To test for the former I compare estimates using both contemporaneous measures of maternal speech with estimates using a lagged measure as a proxy for the contemporaneous measure.

4.4.1. Results: simultaneity tests

I first test for simultaneity bias in Z by comparing contemporaneous measures of home language, Z_{it}, with estimates using lagged measures, $Z_{i,t-1}$ and $Z_{i,t-2}$ at 36 months. I then test whether child language at 36 months is a significant predictor of the home language environment when lagged estimates of the home language environment at 14 and 24 months are included. Columns 1–3 of Table A4 in the Appendix compare a contemporaneous measure of the home language environment with lagged measures. While the 1-period lag registers the best fit, the coefficients in Eqs. (1)–(3) are not statistically different from each other ($F = 0.25$, $p > F = 0.7763$). Column 4 includes all measures of the home environment and column 5 adds a full set of covariates. In both cases, once the lagged measures are accounted for, the contemporaneous measure is not significantly different from zero. Taken together, these results suggest measures of the home language environment record some persistent level of inputs that does not vary greatly over time. Table A5 shows results from a regression of home language factor score at 36 months on child language and lagged versions of home language. Columns 4 and 5 indicate that once prior measures of the home language environment are included in the regression, the child's language score is not a significant predictor.

5. Analysis

5.1. Estimating differential growth rates by race and SES

To test for differential growth rates I estimate a random effects model, including controls for family background characteristics, SES and a measure of the home language environment. Interactions between race and the home language environment and interactions between age, race and SES test for differential returns to parental inputs. To make use of all possible data, I include all observations with valid test scores, inputs and sociodemographic controls in any wave, and include a set of dummies indicating each wave the child has non-missing data. The full specification is:

$$y_{it} = \beta_0 + \beta_1 Age_{it} + \boldsymbol{\beta_2}\mathbf{x_i} + \beta_3 z_{it} + \beta_4(Black_i * Age_{it}) + \boldsymbol{\beta_5}(\tilde{\mathbf{x}}_\mathbf{i} * Black_i) + \beta_6(Black * z_{it}) + \boldsymbol{\beta_7}(\tilde{\mathbf{x}}_\mathbf{i} * Age_{it}) + \boldsymbol{\beta_8}(\tilde{\mathbf{x}}_\mathbf{i} * Black_i * Age_{it}) + \upsilon_i + \epsilon_{it}. \quad (5)$$

where y_{it} is a measure of child language at time t and $\mathbf{x_i}$ is a $M \times 1$ vector of time-invariant background characteristics, including: gender (male), the family's poverty ratio, mother's education (less than high school, high school-GED or more than high school), an indicator of light birthweight, whether the child was first born, the number of siblings age 0–5 living in the home, the mother's age at first birth (centered) and its squared term, a measure of maternal depression on a scale of 0–5, dummies indicating if the Head Start program is a center or a home based program (mixed is the reference group), and a dummy indicating if the family was randomized into the treatment group. $\boldsymbol{\beta_2}$ is a corresponding $1\times M$ coefficient vector. $\tilde{\mathbf{x}}_\mathbf{i}$ is a subset of $\mathbf{x_i}$ containing measures of SES; in the main specification $\tilde{\mathbf{x}}_\mathbf{i}$ includes mother's highest grade completed and the family's poverty ratio and z_{it} is the time-varying home language factor score. Age_{it} is age in months-14 at time t, thus coefficients on the main effects are interpreted as intercepts at 14 months; a quadratic in age is also included in the main effects, but not in the interactions. υ_i is an individual specific effect and ϵ_{it} is an idiosyncratic error term that varies over time and individuals. Because y_{it} is a Z-score and not a cumulative measure, the coefficients are not interpreted as a growth rate, but in terms of relative growth across groups.

Recent studies find differential effects of the home environment across race on early cognitive development. Duncan et al. (1994) find that only the home learning environment is a significant moderating factor in the relationship between SES, race and cognitive development. Farkas and Beron (2004) show two relevant findings. First, that the effect of SES on vocabulary occurs prior to 36 months, and second, that Black respondents receive a significantly lower return to SES than do Whites before 36 months, and that after 36 months high SES Black respondents receive a higher return to SES than do their White peers. These authors create a composite measure of SES using a factor score of mother's highest grade completed, grandmother's highest grade completed, and share of years in poverty. I separate measures, allowing for separate effects of wealth and education on child outcomes. I test three specific hypotheses:

$$H_{0,1} : \beta_4 = 0; \quad H_{0,2} : \beta_6 = 0; \quad H_{0,3} : \boldsymbol{\beta_8} = 0.$$

First, conditional on a full set of time invariant family background characteristics and a time varying measure of the home language environment, $H_{0,1}$ tests whether the language growth rate is different for Black and White children. Second, $H_{0,2}$ tests for interaction effects between the time-varying home language factor score and race, testing whether the effects of the home language environment accrue differently for Black and White children. Third, $H_{0,3}$ tests for differences in language growth by SES, allowing for comparisons of the effects of maternal education and wealth on language development within and across race.

5.1.1. Results: differences in growth patterns across race and SES

Table 4 shows results from estimating Eq. (5) above. The coefficients of interest are interactions between Age and Black, between Age, Black and variables indicating SES (mother's highest grade completed and the family's poverty ratio), and between Black the home language factor score. The dependent variable is the child's language factor score normalized to a Z-score, thus the coefficients are in standard deviations; standard errors are clustered on the individual. Column 1 includes all controls plus an interaction between Age and Black. Results indicate that, conditional on background characteristics, White children gain 0.024*sd* on their Black peers per month. Adding the home language factor score and an interaction with Black in column 2 does not affect the coefficient on Age $*$ Black. The interaction term between Black and home language indicates that a one standard deviation increase in the measured home language environment is associated with a 0.283*sd* increase in child language for White children, but only a 0.185*sd* increase for Black children; a difference of 0.098*sd*.

Lastly, column 3 adds interactions between Age, Black and SES; this is the full specification in Eq. (5). Results yield four substantive findings. First, the interaction between Age and Black is now small and not statistically different from zero, indicating that interactions between race and maternal education over time play an important role in the emergence of the race gap in language acquisition. Second, the coefficient on the interaction between Black and the home language factor score is relatively unchanged. Third, the positive coefficient on the interaction between Black and HGC > 12 indicates that Black children with mothers who have some schooling beyond high school perform better than White children with mothers of the same education level at baseline. Fourth, the coefficient on the interaction between Black, Age and HGC > 12 shows that White children from this group gain 0.032*sd* per month on Black students in the same group. This finding supports Farkas and Beron's (2004) claim that the effect of SES on child development occurs before 36 months for White children, which they see in their data. Combining the coefficients on Age $*$ (HGC > 12), 0.031, and Black $*$ Age $*$ (HGC > 12), -0.032, indicates no significant gain in language growth associated with having more educated mothers for Black children between ages 14 and 36 months.

5.2. Robustness tests 1

I next conduct a series of robustness checks to ensure that findings are not sensitive to alternative model specifications or variable definitions. Column 1 of Table 5 re-estimates the main specification (column 3 of Table 4), adding interactions between Black, Age and *all* time-invariant control variables (this is replacing $\tilde{\mathbf{x}}_\mathbf{i} \subseteq \mathbf{x_i}$ with $\tilde{\mathbf{x}}_\mathbf{i} = \mathbf{x_i}$). The results between the two specifications are similar, especially with respect to the interaction terms between Age, Black and SES. None of the additional interaction terms between Black, Age and the covariates are significant at $p < 0.05$; for brevity these terms and main effects are suppressed from the table. Column 2 re-estimates this specification using fixed effects, differencing out time-invariant heterogeneity. The results vary little from column 2. As expected, the coefficient on the Home Language factor score is small and not significantly different from zero, indicating that this variable measures some largely persistent level of the home language environment.

I next re-estimate the main specification using alternate versions of the home language environment (z_{it}), to ensure that results are not the product of any specific measure of parental speech. Column 3 replaces the home language factor score with

Table 4
Dependent variable is child language factor score, panel regression form three waves. Main specification.

	(1)		(2)		(3)	
Black	0.062	(0.062)	0.163***	(0.061)	−0.080	(0.132)
Boy	−0.276***	(0.042)	−0.267***	(0.040)	−0.264***	(0.040)
Poverty ratio	0.067	(0.045)	0.046	(0.043)	−0.173**	(0.085)
HGC < 12	−0.100	(0.063)	−0.056	(0.060)	−0.012	(0.101)
HGC > 12	0.225***	(0.057)	0.191***	(0.055)	−0.142	(0.094)
Light	−0.469***	(0.102)	−0.468***	(0.096)	−0.478***	(0.095)
Firstborn	−0.099	(0.061)	−0.116**	(0.057)	−0.110*	(0.057)
Siblings (age 0–5)	−0.105***	(0.032)	−0.089***	(0.030)	−0.087***	(0.030)
Age 1st birth (centered)	−0.024***	(0.007)	−0.030***	(0.007)	−0.030***	(0.007)
Age 1st birth^2	0.001*	(0.001)	0.001**	(0.001)	0.001**	(0.001)
Live w/husband	0.027	(0.063)	0.021	(0.060)	0.042	(0.061)
Depression	−0.003	(0.029)	0.016	(0.028)	0.019	(0.028)
Home based	0.062	(0.049)	0.005	(0.047)	0.005	(0.048)
Center based	−0.067	(0.057)	−0.043	(0.055)	−0.051	(0.056)
Treatment group	0.085**	(0.042)	0.074*	(0.040)	0.069*	(0.040)
Age-14 (months)	0.023 ***	(0.006)	0.024***	(0.006)	0.007	(0.007)
Age * Black	−0.024***	(0.004)	−0.025***	(0.004)	−0.008	(0.009)
Home language			0.283***	(0.037)	0.281***	(0.037)
Black * Home language			−0.098**	(0.044)	−0.105**	(0.045)
Black * Poverty ratio					0.205	(0.130)
Age * Poverty ratio					0.012**	(0.005)
Black * Age * Poverty ratio					−0.005	(0.009)
Black * HGC < 12					−0.007	(0.139)
Black * HGC > 12					0.325**	(0.148)
Age * HGC < 12					−0.001	(0.006)
Age * HGC > 12					0.031***	(0.006)
Black * Age * HGC < 12					−0.007	(0.009)
Black * Age * HGC > 12					−0.032***	(0.009)
N	2881		2881		2881	
Clusters	1458		1458		1458	

Dependent variable is child language factor score normalized to a *Z*-score. Home Language is a factor score from CFA on items from HOME and 3-bag evaluation at each wave, also normalized to *Z*-scores. All models include *age*^{2} and dummies indicating in which waves the respondent has valid data. Missing values for poverty ratio are coded to zero and a dummy variable is set equal to 1 if missing.
Standard errors clustered on individuals in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

separate factor scores for the 3-bag evaluation and the HOME score. Column 4 includes a measure of maternal language ability using the Woodcock–Johnson evaluation (Woodcock, 1987); the sample used is slightly smaller as 409 observations lacked this measure. The addition of the Woodcock score has little effect on results, likely because the Woodcock is a proxy for maternal education which is already accounted for in the estimation. Column 5 substitutes a lagged measure of the home language environment for the contemporaneous measure. This estimation uses only those observations with a valid lagged home language score, thus only respondents in waves 2 and 3 of the survey are present. The coefficients on the Lagged home language score and interaction with Black are similar to those in the original specification. The main effect for home language is slightly larger, by 0.036*sd*, and the negative interaction with Black is also smaller (more negative) by 0.067. The coefficients on interactions between Age and SES variables are robust to this specification as well. Lastly, column 6 replaces the home language score in each wave with its mean over the three waves, and re-estimates the full specification in column 1 including all interactions. This increases the sample by including observations that were dropped from because they were missing the HOME evaluation. The purpose is to ensure that the main result is not biased by omitting these observations.

5.3. Robustness tests 2

In Table 6 I re-estimate Eq. (5) using all possible combinations of dependent variables as opposed to the factor score method. This is equivalent to subscripting y_{it} in Eq. (5) with $y_{it,j}$ for each test *j* at age *t*. The purpose is to verify that the results are not particular to any one set of tests. That is, it is possible that instead of taking three language outcome measures at 14 and 24 months, researchers could have only taken one. How would this have changed my interpretation of the evolution of the Black–White gap in language? Column 1 of Table 6 re-estimates the main specification in column 3 of Table 4 for reference. Columns 2–10 re-estimate the primary specification using all possible test combinations. While there is substantial comparability across columns 2–7, there is some variation in columns 8–10 – those using MacArthur Production as the 14 month measure. In these specifications, the interaction term between Black and Age is negative and significantly different from zero, and the main effect for Black is positive, although not statistically different from zero. This results from the fact

Table 5
Robustness checks 1 – dependent variable is child language factor score, panel regression from three waves.

	(1)		(2)		(3)		(4)		(5)		(6)	
Black	−0.489*	(0.297)			−0.065	(0.132)	−0.025	(0.160)	−0.395	(0.286)	−0.011	(0.121)
Poverty ratio	−0.174**	(0.085)			−0.173**	(0.086)	−0.199**	(0.101)	0.032	(0.169)	−0.180**	(0.083)
HGC < 12	0.030	(0.112)			−0.005	(0.101)	0.034	(0.122)	−0.219	(0.217)	−0.005	(0.097)
HGC > 12	−0.112	(0.098)			−0.146	(0.094)	−0.155	(0.109)	−0.320	(0.208)	−0.152	(0.093)
Black * Poverty ratio	0.245*	(0.130)			0.206	(0.129)	0.243	(0.149)	0.224	(0.233)	0.163	(0.120)
Black * HGC < 12	−0.109	(0.168)			−0.012	(0.139)	−0.141	(0.167)	0.318	(0.301)	−0.026	(0.131)
Black * HGC > 12	0.318**	(0.157)			0.328**	(0.148)	0.257	(0.177)	0.398	(0.311)	0.291**	(0.142)
Black * age	0.008	(0.019)	−0.008	(0.031)	−0.008	(0.009)	−0.005	(0.011)	0.006	(0.016)	−0.009	(0.008)
Age * poverty ratio	0.012**	(0.005)	0.013	(0.009)	0.012**	(0.005)	0.011*	(0.006)	0.001	(0.009)	0.012**	(0.005)
Age * HGC < 12	−0.005	(0.007)	0.000	(0.011)	−0.002	(0.006)	−0.002	(0.007)	0.009	(0.011)	−0.002	(0.005)
Age * HGC > 12	0.029***	(0.006)	0.035***	(0.009)	0.031***	(0.006)	0.029***	(0.007)	0.039***	(0.011)	0.030***	(0.006)
Black * age * poverty ratio	−0.007	(0.009)	−0.008	(0.015)	−0.004	(0.009)	−0.004	(0.011)	−0.006	(0.014)	−0.001	(0.008)
Black * age * HGC < 12	0.003	(0.010)	−0.004	(0.016)	−0.007	(0.009)	−0.000	(0.010)	−0.021	(0.016)	−0.007	(0.008)
Black * age * HGC > 12	−0.033***	(0.010)	−0.040**	(0.016)	−0.032***	(0.009)	−0.031***	(0.011)	−0.034**	(0.017)	−0.029***	(0.009)
Home language	0.286***	(0.038)	0.081	(0.084)			0.232***	(0.043)				
Black * Home language	−0.114**	(0.046)	−0.020	(0.102)			−0.070	(0.052)				
HOME language factor					0.174***	(0.037)						
3-Bag factor					0.163***	(0.032)						
Black * HOME language factor					−0.047	(0.044)						
Black * 3-bag factor					−0.080*	(0.042)						
Woodcock–Johnson							0.034	(0.047)				
Age * Woodcock							0.007**	(0.003)				
Black * Woodcock							0.044	(0.074)				
Black * age * Woodcock							−0.002	(0.004)				
Lagged home language									0.317***	(0.045)		
Black * Lagged home language									−0.172***	(0.057)		
Mean home language factor											0.410***	(0.047)
Black * Mean home language											−0.153***	(0.056)
All controls	X		X		X		X		X		X	
Age * Black * All controls	X		X									
N	2881		2881		2881		2258		1473		3257	
Clusters	1458		1458		1458		1049		965		1484	

Dependent variable is child language factor score normalized to a *Z*-score. All Controls includes all control variables in column 1 of Table 4. *Age * Black * All Controls* indicates all main effects and interactions between Age, Black and all covariates from column 1 of Table 4. HOME and 3-bag factors are factor scores created separately using items from each evaluation. Woodcock is mother's Woodcock–Johnson vocabulary score. Home language lag is the Home Language score from the previous wave. Column 6 replaces Home Language in each wave with the average over the three waves.
Standard errors clustered on individuals in parentheses.

* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

Table 6
Robustness checks 2 – dependent variable is child language at each wave.

Measure 14 Measure 24 Measure 36	Factor Factor Factor (1)	Gesture Sentence PPVT (2)	Gesture Prod PPVT (3)	Gesture Comp PPVT (4)	Comp Sentence PPVT (5)	Comp Prod PPVT (6)	Comp Comp PPVT (7)	Prod Sentence PPVT (8)	Prod Prod PPVT (9)	Prod Comp PPVT (10)
Black	−0.080 (0.132)	−0.107 (0.126)	−0.071 (0.130)	−0.124 (0.126)	−0.153 (0.129)	−0.118 (0.135)	−0.163 (0.128)	0.117 (0.128)	0.165 (0.133)	0.113 (0.127)
Home language	0.281*** (0.037)	0.276*** (0.038)	0.260*** (0.041)	0.265*** (0.040)	0.267*** (0.036)	0.248*** (0.040)	0.261*** (0.038)	0.252*** (0.033)	0.244*** (0.034)	0.251*** (0.034)
Black * Home language	−0.105** (0.045)	−0.110** (0.045)	−0.109** (0.049)	−0.096** (0.049)	−0.111** (0.043)	−0.108** (0.047)	−0.098** (0.047)	−0.087** (0.042)	−0.092** (0.043)	−0.079* (0.044)
Black * Poverty ratio	0.205 (0.130)	0.255* (0.140)	0.203 (0.139)	0.216* (0.128)	0.213* (0.125)	0.162 (0.127)	0.174 (0.116)	0.136 (0.136)	0.078 (0.138)	0.091 (0.126)
Black * HGC < 12	−0.007 (0.139)	0.020 (0.136)	0.050 (0.139)	0.075 (0.137)	0.071 (0.141)	0.101 (0.145)	0.123 (0.139)	−0.205 (0.137)	−0.181 (0.139)	−0.156 (0.135)
Black * HGC > 12	0.325** (0.148)	0.216 (0.147)	0.307** (0.151)	0.249* (0.147)	0.295** (0.149)	0.387** (0.151)	0.325** (0.146)	0.171 (0.150)	0.258* (0.150)	0.199 (0.147)
Age * Black	−0.008 (0.009)	−0.006 (0.009)	−0.006 (0.009)	−0.006 (0.009)	−0.004 (0.009)	−0.005 (0.009)	−0.006 (0.009)	−0.018** (0.008)	−0.019** (0.008)	−0.019** (0.008)
Age * Poverty ratio	0.012** (0.005)	0.013** (0.006)	0.013** (0.006)	0.011** (0.006)	0.011** (0.005)	0.011** (0.005)	0.010* (0.005)	0.009** (0.005)	0.010** (0.005)	0.008* (0.005)
Age * HGC < 12	−0.001 (0.006)	0.001 (0.006)	0.002 (0.006)	0.003 (0.006)	−0.003 (0.006)	−0.002 (0.006)	−0.001 (0.006)	−0.004 (0.005)	−0.003 (0.005)	−0.001 (0.005)
Age * HGC > 12	0.031*** (0.006)	0.033*** (0.006)	0.033*** (0.006)	0.032*** (0.006)	0.027*** (0.006)	0.027*** (0.006)	0.025*** (0.006)	0.032*** (0.005)	0.032*** (0.005)	0.030*** (0.005)
Black * Age * Poverty ratio	−0.005 (0.009)	−0.008 (0.009)	−0.006 (0.009)	−0.005 (0.010)	−0.004 (0.008)	−0.002 (0.008)	−0.001 (0.008)	−0.002 (0.009)	−0.000 (0.009)	0.001 (0.009)
Black * Age * HGC < 12	−0.007 (0.009)	−0.009 (0.009)	−0.008 (0.009)	−0.011 (0.009)	−0. 012 (0.009)	−0.011 (0.009)	−0.014 (0.009)	0.003 (0.008)	0.004 (0.008)	0.000 (0.008)
Black * Age * HGC > 12	−0.032*** (0.009)	−0. 028*** (0.010)	−0.029*** (0.010)	−0. 028*** (0.009)	−0.031*** (0.010)	−0. 033*** (0.010)	−0.031*** (0.010)	−0. 024*** (0.009)	−0.026*** (0.009)	−0. 024*** (0.009)
Controls and main effects	X	X	X	X	X	X	X	X	X	X
N	2881	2881	2881	2881	2881	2881	2881	2881	2881	2881
Clusters	1458	1458	1458	1458	1458	1458	1458	1458	1458	1458

Controls and main effects include all control variables in column 1 of Table 4 and main effects for Age, Poverty ratio, HGC, and Home language.
Standard errors clustered on individuals in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

that the Black–White gap in language production in smaller than the gap in other measures, and in fact favors Black children at 14 months. Despite subtile differences across measures, these results support the overall conclusion that Black children appear to benefit less from parental education and from the home language environment than do White children in terms of language acquisition, and that these differences pose significant contributions to the emergence of the Black–White gap in language skill in the first few years.

5.4. Decomposition

To further analyze this hypothesis, I decompose the share of the observed language gap in each wave (Oaxaca, 1973).[5] This determines what share of the total language gap is due to differences in endowments (each of the X's), the share due to Black–White differences in returns to these endowments (β's), and the share due to interactions between these. Using notation from Jann (2008), let R_{W-B} be the mean (White–Black) gap in measured language at each wave, where

$$R_{W-B} = E[Y_W] - E[Y_B] = E[X'_W]\beta_W - E[X'_B]\beta_B. \tag{6}$$

Terms $E[X'_W]$ and β_W can be re-written as:

$$E[X'_W] = E[X'_B] + (E[X'_W] - E[X'_B]), \tag{7a}$$

$$\beta_W = \beta_B + (\beta_W - \beta_B), \tag{7b}$$

where $(E[X'_W] - E[X'_B])$ is the difference between Black and White children in the covariates ($\mathbf{x}_i'$s and z_{it}), and $(\beta_W - \beta_B)$ is the differential return White children receive (in terms of language development) compared to Black children. Plugging (7a) and (7b) into (6) and rearranging yields:

$$R_{W-B} = \underbrace{(E[X'_W] - E[X'_B])\beta_B}_{\text{endowments}} + \underbrace{E[X'_B](\beta_W - \beta_B)}_{\text{coefficients}} + \underbrace{(E[X'_W] - E[X'_B])(\beta_W - \beta_B)}_{\text{interactions}}. \tag{8}$$

Interacting $(E[X'_W] - E[X'_B])$ with β_B in the first term shows how much Black children would gain on average if they were moved to the same levels of $\mathbf{x_i}$ and z_{it} as White children, at their current level of β_B. The second term in Eq. (8) shows how much Black children would gain on average if they were moved to the same coefficient levels as White children, given their current levels of the covariates. The third term, $(E[X'_W] - E[X'_B])(\beta_W - \beta_B)$, is an interaction between these two effects showing the benefit Black children would gain if they were moved to White levels of both covariates and coefficients.

Table 7 shows results from the decomposition. Paying attention first to the home language factor score, at 14 and 24 months Black children would gain about 0.13*sd* in measured language acquisition on average if they were moved to the average White levels of the home language environment, and 0.096*sd* at 36 months. Yet, the decomposition shows a negative effect in the case that the average Black home language environment is held constant and Black children are moved to the same coefficient levels for the home language environment as White children at each age. The interaction effect between the two is positive. Turning next to maternal education, I find that at 14 months Black children would be worse off (by −0.080*sd*) if they were moved to the same coefficient levels as White children in terms of mothers' having earned education beyond high school. At 36 months, they would be better off by a roughly equal amount (0.086*sd*). At the bottom of Table 7, the Totals indicate that at 14 months Black children received higher returns to the covariates than White children, at 24 months the language gap was largely explained by differences in the covariates, and that at 36 months the gap was attributable to both lower levels of and returns to the covariates.

5.5. Discussion

While the longitudinal analysis suggests a comparatively steeper language learning gradient according to maternal education for White children, the decomposition above suggests that returns to maternal education benefit Black children at 14 months and White children at 36 months. I now turn to a possible explanation for this. Feldman et al. (2000) evaluate a large and diverse sample of 1 and 2 year-old children using the MacArthur CDI. Similar to results here, the authors find lower scores for both White children and for high SES groups (identified by maternal education and private health insurance respectively) on measures of phrases understood, vocabulary comprehension and vocabulary production, but not for gesturing, at age 1. In a norming study of the MacArthur CDI, Fenson et al. (1993) find similar results for vocabulary comprehension at age 1 for mothers with less than a high school degree, as did Reznick (1990) using an earlier version of the MacArthur. Feldman et al. (2000) claim that on subsections requiring "considerable subjective interpretation," poorer and less educated parents, who were underrepresented in the CDI norming sample, seemingly *overestimate* their child's language compared with more educated and wealthier parents, especially at age 1. They identify the inherently more subjective evaluation of language at the earliest stages of acquisition as a possible explanation for this. They conclude that by age 2 parents make accurate evaluations of productive vocabulary and word combinations, but that evaluations of irregular word forms may be less accurate.

While Feldman et al.'s (2000) disclaimer, supported by the data here, might explain differences across maternal education, they do little to explain differences across race after controlling for SES, and do little to explain interactions between

[5] See also Jones and Kelley (1984) and Winsborough and Dickenson (1971).

Table 7
Decomposition of the Black–White gap in PPVT, 36 months.

	14 months		24 months		36 months	
Differential						
$E[Y_W]$	−0.089**	(0.042)	0.142***	(0.046)	0.246***	(0.043)
$E[Y_B]$	−0.026	(0.047)	−0.182***	(0.044)	−0.277***	(0.046)
$E[Y_W] - E[Y_B]$	−0.062	(0.063)	0.325***	(0.064)	0.524***	(0.064)
Endowments						
Boy	0.007	(0.008)	0.007	(0.006)	0.004	(0.005)
Poverty ratio	0.002	(0.015)	0.026*	(0.015)	0.022	(0.014)
HGC < 12	−0.014	(0.020)	0.023	(0.019)	0.006	(0.015)
HGC > 12	0.013	(0.010)	0.012	(0.012)	0.006	(0.007)
Light birthweight	0.006	(0.012)	0.004	(0.004)	0.010	(0.009)
Firstborn	0.001	(0.016)	0.015	(0.013)	0.021*	(0.013)
Siblings (age 0–5)	−0.004	(0.007)	−0.010	(0.008)	−0.014	(0.010)
Age 1st birth	−0.031	(0.020)	−0.034**	(0.016)	−0.012	(0.011)
Live w/husband	−0.005	(0.057)	0.111**	(0.044)	0.011	(0.049)
Depression	−0.026	(0.033)	−0.091**	(0.036)	0.007	(0.033)
Treatment group	0.001	(0.005)	−0.003	(0.006)	0.002	(0.004)
Age	−0.005	(0.013)	0.059***	(0.020)	0.000	(0.006)
Home language	0.137***	(0.031)	0.129***	(0.025)	0.096***	(0.026)
Coefficients						
Boy	−0.036	(0.061)	−0.098	(0.065)	−0.062	(0.062)
Poverty ratio	−0.094	(0.057)	−0.028	(0.053)	−0.053	(0.056)
HGC < 12	0.042	(0.074)	0.003	(0.081)	0.029	(0.077)
HGC > 12	−0.080**	(0.041)	0.015	(0.039)	0.086**	(0.041)
Light birthweight	0.023	(0.016)	−0.033	(0.021)	0.011	(0.019)
Firstborn	−0.072	(0.121)	0.136	(0.119)	0.026	(0.107)
Siblings (age 0–5)	0.010	(0.099)	0.099	(0.101)	−0.026	(0.092)
Age 1st birth	−0.164	(0.280)	−0.061	(0.288)	0.079	(0.272)
Live w/husband	−0.004	(0.017)	−0.025*	(0.015)	0.006	(0.017)
Depression	−0.285	(0.223)	−0.172	(0.230)	0.164	(0.228)
Treatment group	−0.097*	(0.057)	−0.064	(0.063)	−0.045	(0.062)
Age	0.602	(0.758)	−1.035	(1.101)	2.007	(1.494)
Home language	−0.045*	(0.026)	−0.026	(0.020)	−0.068**	(0.027)
Interactions						
Boy	0.002	(0.004)	0.007	(0.008)	0.004	(0.006)
Poverty ratio	−0.032	(0.021)	−0.011	(0.021)	−0.017	(0.019)
HGC < 12	−0.015	(0.026)	−0.001	(0.029)	−0.008	(0.021)
HGC > 12	−0.019	(0.013)	0.006	(0.015)	0.016	(0.013)
Light birthweight	−0.003	(0.006)	0.010	(0.010)	−0.003	(0.006)
Firstborn	0.012	(0.021)	−0.021	(0.019)	−0.003	(0.013)
Siblings (age 0–5)	0.001	(0.010)	0.008	(0.009)	−0.002	(0.007)
Age 1st birth	−0.014	(0.024)	−0.004	(0.020)	0.004	(0.015)
Live w/husband	−0.015	(0.064)	−0.096*	(0.057)	0.020	(0.056)
Depression	0.055	(0.044)	0.039	(0.052)	−0.030	(0.042)
Treatment group	−0.001	(0.006)	0.002	(0.004)	−0.001	(0.003)
Age	−0.001	(0.003)	−0.013	(0.014)	0.011	(0.010)
Home language	0.085*	(0.048)	0.053	(0.040)	0.122***	(0.045)
Totals						
Endowments	0.079	(0.066)	0.225***	(0.049)	0.147***	(0.051)
Coefficients	−0.201***	(0.075)	0.092	(0.073)	0.258***	(0.075)
Interactions	0.060	(0.076)	0.008	(0.061)	0.119*	(0.065)
N	998		936		947	

Dependent variable is child language at each wave; White and Black refer to the reference group. Not shown are dummies for home or center based Head Start and the constant term.
Robust standard errors are in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

race and SES measures. A robust literature chronicles differences between Black and White speech[6]; it may be the case that these differences are not appropriately reflected in the language measures presented here, especially at early ages. To further analyze this issue, I exploit a unique aspect of this data: at 24 months there is an "objective" measure of child language, from the

[6] See Labov (1972), Clopper and Pisoni (2004), Bailey and Thomas (1998), Rickford and Rafal (1996), Washington and Craig (2002), and Wolfram (1971, 1991).

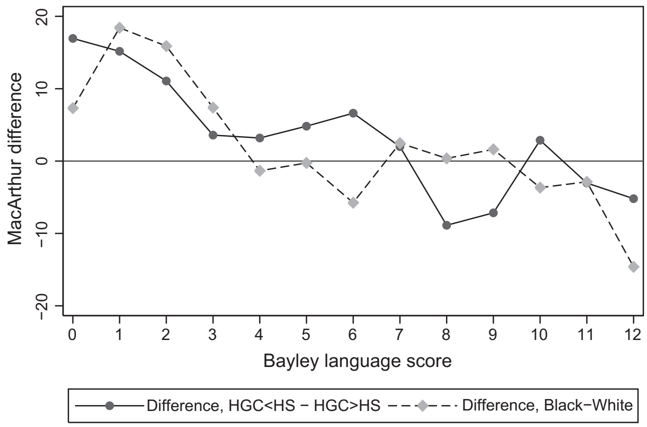

Fig. 4. Differences in MacArthur score by Bayley language score at 24 months. Figure plots the difference in maternal report of productive vocabulary at 24 months (solid: HGC < 12–HGC > 12; dashed: Black–White), by child's Bayley language items score. MacArthur difference is a word count from maternal report of 100 possible items. Bayley Language score is the sum of 12 language items from the Bayley test that resulted from factor analysis conducted by the principal investigators; items are listed in the appendix (see Love et al. (2001) for further information).

Bayley language items, and a "subjective" parental report, from the MacArthur production measure (objective here refers to a a measure that is consistent across interviewers, as opposed to varying by parents). Fig. 4 plots differences in MacArthur production reported by parents, by Bayley language score outcomes reported by interviewers (scores range from 0 to 12; specific items are listed in the Appendix). The first (solid) line plots differences by maternal education (HGC < 12–HGC > 12) for each value of Bayley language; the second (dashed) line repeats this exercise by race. That is, Fig. 4 plots differences across race and maternal education in "subjective" measures, by the "objective" scores. Results indicate that for children who scored poorly on the "objective" measure of child language, less educated parents (HGC < 12) consistently report that their children say more words than do more educated mothers (HGC > 12). Repeating this procedure by race (Black–White) yields similar results; for children who scored poorly on the "objective" measure, Black mothers report more language production than their White peers. Repeating this exercise using the MacArthur Sentence Completion measure at 24 month reveals a somewhat similar pattern, although it is far less pronounced and group differences are not significantly different from zero, thus results are not shown.

There are several possible explanations for these results. On one hand, it might be the case that low SES parents overestimate their children's language production at early ages, which would understate White–Black language gaps at 14 and 24 months in the present study. Conversely, it might be the case that "objective" language measures, such as the Bayley language items measure and the PPVT, do not accurately measure language for African–American respondents, which would overstate the White–Black gap at 24 months (for the Bayley score only) and at 36 months. From the former case, a third possibility arises. Language theory implies that there should be some "optimal" level of child directed speech given a child's current level of language ability. Thus, parents who overestimate their child's language ability will speak to their children using language that is too complex. Analogously, parents who underestimate their child's language ability might under-stimulate using language that is too simplistic. This fits with evidence that parents adjust child-directed speech in terms of both quantity and quality as child language develops (Rowe et al., 2005), and evidence that more educated mothers both spend more time (Guryan et al., 2008), and tailor their child interactions more appropriately to the child's developmental stage (Kalil et al., forthcoming), than do less educated mothers.

Results from this analysis, though not conclusive, shed some light on this relationship and suggest that while measurement might explain some of the observed differential across race (and possibly SES), it is unlikely that it accounts for the total differential. Results presented in Table 6 suggest that interaction effects between race and maternal education are robust to various language measures, and results in Table 7 indicate that some of this difference is attributable to differential returns across race. Further research and more data are needed to entirely disentangle these discrepancies.

6. Conclusion

In this study I estimate comparative growth models for Black and White children using data from the Early Head Start Research and Evaluation study with several measures of language skill and the home language environment at 14, 24, and 36 months. I test whether a diverse set of controls for background characteristics, income, SES and for the home language environment – a proxy for the amount, type and quality of speech children hear in the home – can account for differences in language development over the first 3 years of life, and whether the effects of these measures accrue differently for Black and White children over this time period.

Results indicate several significant findings. First, the Black–White gap in in the home language environment, measured by a principal factor score from linguistic variables from the HOME and 3-bag evaluations of the child's learning environment, is over a half a standard deviation at 14 months and remains relatively constant over the 22 month observation period. Second, although at 14 months Black children slightly outperform their White peers, by 24 months a significant gap in language emerges which widens by 36 months; these gaps persist despite a robust set of controls for demographic characteristics, family circumstances and the home language environment. Third, persistent negative coefficients on interactions between age, race and maternal education, and between race and a time-varying measure of the home language environment confirm Farkas and Beron's (2004) hypothesis that returns to SES and measures of the home language environment accrue differently with respect to race over these ages. Lastly, a decomposition of the gap at each wave indicates that while at 24 months differences in endowments of the covariates explain much of the measured language gap, at 36 months only half of the race gap is due to differences in endowments of the covariates, and half is attributable to differences in returns to these endowments, in particular to differential returns to maternal education and a measure of the home language environment. Several robustness checks verify these results, including: alternate measures of the home language environment, a test for simultaneity in maternal and child language measures using lagged inputs and outputs as instruments for contemporaneous measures, and re-estimating the main specification using all possible combinations of child language measures at each age as opposed to the factor score measure to ensure that the results are not dependent on any one language measure.

I add to the literature by showing the importance of interactions between race and maternal education, and between race and the home language environment, in the emergence of racial skill gaps at early ages. In addition, I address methodological concerns in estimating group differences in language acquisition with respect to both model misspecification and bias from measurement error. I also address the subjective nature of parental reported child language, and discuss the significance of linguistic differences between Black and White English speakers in terms of measuring child language growth. These findings shed light on a previously undocumented aspect of differences in language learning across race and highlight the existence of unobserved heterogeneity across race in the home environment that is not accounted for by traditional measures of parental education, the home environment, SES or income. Moreover, these findings emphasize the importance of the home environment on language development and suggest that interventions targeting early language learning should address the home learning environment in addition to providing location-based services.

Acknowledgement

Thanks to Jeff Grogger, Kerwin Charles, Amy Claessens, Meredith Rowe, Rebecca Ryan and two anonymous referees for suggestions. Thanks also the University of Chicago's Committee on Education and the Institute of Education Sciences Predoctoral Interdisciplinary Research Training Programs for funding support. Opinions reflect those of the author and do not necessarily reflect those of the granting agencies.

Appendix

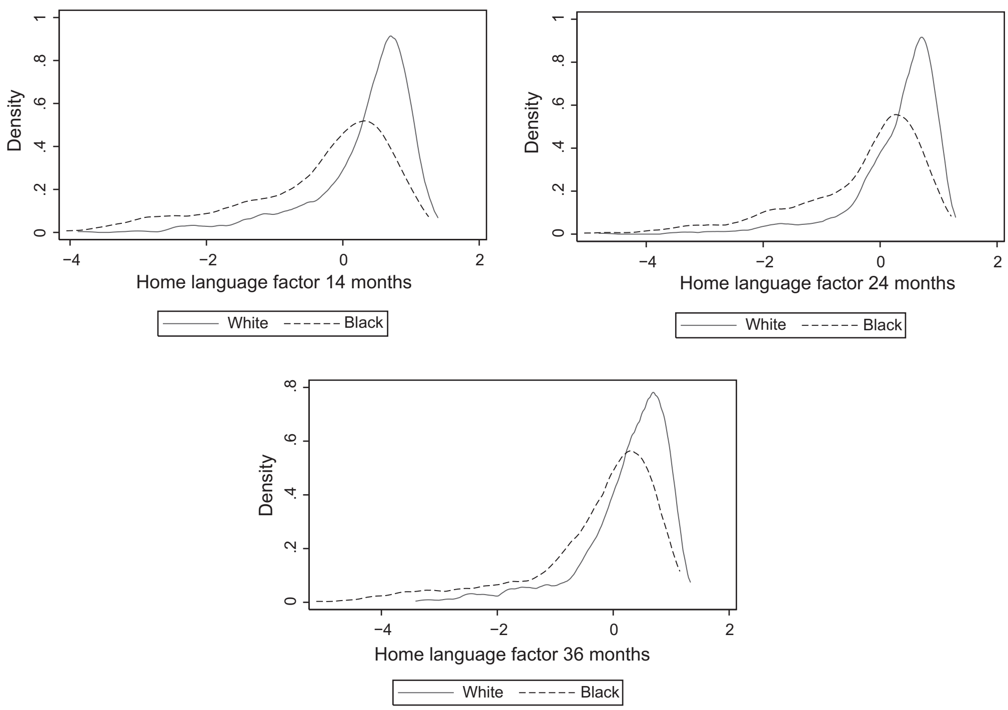

Fig. A1. Density plots of Home Language factor scores, by race and wave. Distribution of Child Language factor score at 14, 24, and 36 months from working sample.

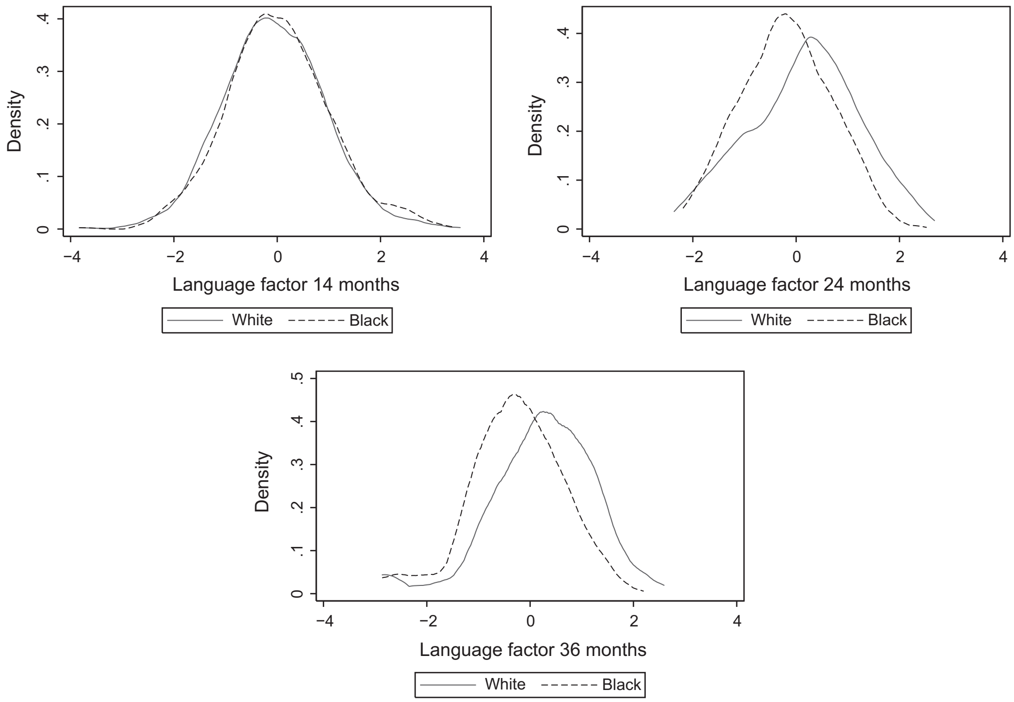

Fig. A2. Density plots of child language factor scores, by race and wave. Distribution of home language factor score at 14, 24, and 36 months from working sample.

Table A1
Summary data, means and standard deviations by data availability.

	14 months		24 months		36 months		With ⩾1 lag		All waves	
	Mean	*sd*	Mean	*sd*	Mean	*sd*	Mean	*sd*	Mean	*sd*
Black	0.47	(0.50)	0.43	(0.50)	0.43	(0.49)	0.42	(0.49)	0.37	(0.48)
Boy	0.51	(0.50)	0.52	(0.50)	0.50	(0.50)	0.51	(0.50)	0.52	(0.50)
Poverty ratio	0.37	(0.48)	0.37	(0.48)	0.38	(0.49)	0.37	(0.48)	0.32	(0.47)
Poverty missing	0.29	(0.45)	0.30	(0.46)	0.27	(0.45)	0.30	(0.46)	0.33	(0.47)
HGC < 12	0.61	(0.49)	0.62	(0.54)	0.60	(0.50)	0.62	(0.53)	0.67	(0.51)
HGC > 12	0.16	(0.37)	0.15	(0.36)	0.16	(0.37)	0.16	(0.37)	0.15	(0.36)
Light birth	0.05	(0.23)	0.06	(0.24)	0.06	(0.23)	0.06	(0.23)	0.05	(0.23)
Firstborn	0.66	(0.47)	0.64	(0.48)	0.64	(0.48)	0.64	(0.48)	0.64	(0.48)
Siblings age 0–5	1.12	(0.77)	1.16	(0.81)	1.14	(0.80)	1.16	(0.79)	1.09	(0.75)
Age 1st birth	22.09	(5.57)	22.12	(5.54)	22.08	(5.49)	22.20	(5.55)	22.55	(5.66)
Live with husband	0.23	(0.42)	0.25	(0.43)	0.23	(0.42)	0.24	(0.43)	0.28	(0.45)
Depression	2.55	(1.19)	2.53	(1.24)	2.57	(1.24)	2.52	(1.22)	2.33	(1.24)
Home based	0.40	(0.49)	0.43	(0.50)	0.41	(0.49)	0.41	(0.49)	0.42	(0.49)
Center based	0.25	(0.43)	0.24	(0.42)	0.27	(0.44)	0.26	(0.44)	0.30	(0.46)
Treatment	0.50	(0.50)	0.53	(0.50)	0.53	(0.50)	0.53	(0.50)	0.51	(0.50)
14 month sample	1.00	(0.00)	0.65	(0.48)	0.66	(0.48)	0.73	(0.45)	1.00	(0.00)
24 month sample	0.61	(0.49)	1.00	(0.00)	0.65	(0.48)	0.88	(0.33)	1.00	(0.00)
36 month sample	0.62	(0.48)	0.66	(0.47)	1.00	(0.00)	0.77	(0.42)	1.00	(0.00)
N	998		936		947		965		426	

Sample from EHSRE includes all Black and White native English speakers with complete HOME and 3-bag evaluations in at least one wave. Each cross-sectional sample includes all observations with both maternal language and child speech measures in that wave. Those in the lagged version have a valid child language, a home language score and lagged home language at either 24 or 36 months. All observations in the *all waves* sample are necessarily in the other 4 samples.

Table A2
Summary statistics by race and HGC, means and standard deviations.

	HGC < 12				HGC = 12				HGC > 12			
	White		Black		White		Black		White		Black	
Poverty ratio	0.51	(0.41)	0.44	(0.48)	0.67	(0.45)	0.50	(0.48)	0.75	(0.48)	0.70	(0.66)
Light birth	0.06	(0.23)	0.08	(0.27)	0.05	(0.21)	0.06	(0.23)	0.05	(0.21)	0.08	(0.26)
First born	0.67	(0.47)	0.83	(0.38)	0.59	(0.49)	0.58	(0.49)	0.53	(0.50)	0.60	(0.49)
Siblings, age 0–5	1.11	(0.85)	1.05	(0.84)	1.25	(0.70)	1.21	(0.83)	1.21	(0.80)	1.09	(0.76)
Treatment	0.54	(0.50)	0.51	(0.50)	0.53	(0.50)	0.48	(0.50)	0.48	(0.50)	0.56	(0.50)
Age first birth	20.30	(5.24)	18.33	(4.41)	22.80	(4.72)	22.56	(4.85)	25.23	(5.30)	24.31	(5.03)
Live w/husband	0.25	(0.43)	0.02	(0.14)	0.37	(0.48)	0.12	(0.31)	0.46	(0.50)	0.12	(0.33)
Depression (0–5)	3.45	(0.85)	3.71	(0.73)	2.03	(0.98)	2.52	(0.92)	1.54	(1.01)	1.97	(0.86)
n	248		319		292		202		234		163	
N	1458											

Sample from EHSRE includes all Black and White native English speakers with valid test scores and inputs in at least one wave.

Table A3
Cross-correlation table, language factor scores, child language tests, and input factors.

	Home lang. factor, 14	Child lang. factor, 14	Gesture, 14	Comprehension, 14	Production, 14	Home lang. factor, 24	Child lang. factor, 24	Production, 24	Complexity, 24	Bayley, 24	Home lang. factor, 36	PPVT, 36
14 Months												
Home language factor	1.000											
Child language factor	0.199	1.000										
MacArthur gesture	0.170	0.731	1.000									
MacArthur comprehension	0.149	0.871	0.479	1.000								
MacArthur production	0.147	0.787	0.380	0.517	1.000							
24 Months												
Home language factor	0.624	0.136	0.103	0.098	0.102	1.000						
Child language factor	0.273	0.458	0.314	0.377	0.373	0.250	1.000					
MacArthur production	0.212	0.496	0.374	0.424	0.379	0.195	0.882	1.000				
MacArthur complexity	0.205	0.364	0.245	0.287	0.323	0.211	0.834	0.605	1.000			
Bayley language	0.272	0.273	0.204	0.192	0.232	0.252	0.752	0.520	0.460	1.000		
36 Months												
Home language factor	0.498	0.094	0.104	0.073	0.042	0.555	0.146	0.146	0.113	0.173	1.000	
PPVT	0.346	0.212	0.181	0.170	0.147	0.324	0.410	0.343	0.287	0.439	0.346	1.000

Pairwise correlations from $N = 1458$ observations with at least one valid test scores and home language score in a wave. Home language factors refer to factor scores estimated from HOME and 3-bag items, and child language factors refer to scores estimated from all available child language measures at each wave. All tests and factor scores are normalized to *Z*-scores from the full sample. All correlation coefficients are significant at $\alpha = 0.01$.

Table A4
Dependent variable is PPVT (*Z*-score) at 36 months.

	(1)	(2)	(3)	(4)	(5)
Home language, 36 months	0.323***			0.098	0.088
	(0.053)			(0.064)	(0.061)
Home language, 24 months		0.377***		0.203***	0.182**
		(0.058)		(0.077)	(0.074)
Home language, 14 months			0.368***	0.195***	0.195***
			(0.045)	(0.064)	(0.066)
Controls					X
N	426	426	426	426	426
R^2	0.084	0.125	0.119	0.156	0.264

Sample includes respondents with child language and home language factor scores in all three waves. Home language is factor score from CFA on items from HOME and 3-bag evaluation at each wave, normalized to *Z*-scores. Controls include gender, race, poverty ratio, mother's HGC, indicator for light birth, first born, number of siblings living in the home ages 0–5, mother's age at 1st birth and squared term, indicator if living with husband (not necessarily child's father), sum of 5 items indicating mother's depression, dummies for home based or center based Head Start, dummy for assignment to Head Start treatment group.
Robust standard errors in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

Table A5
Dependent variable is Home Language factor (*Z*-score) at 36 months.

	(1)	(2)	(3)	(4)	(5)
PPVT, 36 months	0.259***	0.088**	0.117***	0.059	0.064
	(0.047)	(0.041)	(0.043)	(0.040)	(0.041)
Home language, 24 months		0.513***		0.390***	0.357***
		(0.059)		(0.068)	(0.063)
Home language, 14 months			0.438***	0.218***	0.218***
			(0.059)	(0.063)	(0.062)
Controls					X
N	426	426	426	426	426
R^2	0.084	0.338	0.270	0.369	0.403

Sample includes respondents with child language and home language factor scores in all three waves. Home language is factor score from CFA on items from HOME and 3-bag evaluation at each wave, normalized to *Z*-scores. PPVT is child's score on Peabody Picture Vocabulary Test, normalized to a *Z*-score Controls include gender, race, poverty ratio, mother's HGC, indicator for light birth, first born, number of siblings living in the home ages 0–5, mother's age at 1st birth and squared term, indicator if living with husband (not necessarily child's father), sum of 5 items indicating mother's depression, dummies for home based or center based Head Start, dummy for assignment to Head Start treatment group.
Robust standard errors in parentheses.
* $p < 0.10$
** $p < 0.05$.
*** $p < 0.01$.

Home Observation for Measurement of the Environment (HOME)[7]

HOME Emotional Responsivity, 14 and 24 months:
Parent spontaneously vocalized to child twice
Parent responds verbally to child's verbalizations
Parent tells child name of object or person during visit
Parent spontaneously praises child at least twice
Parent's voice conveys positive feelings toward child
Parent caresses or kisses child at least once
Parent responds positively to praise of child offered by visitor

[7] Love et al. (2001).

HOME Warmth Scale, 36 months:
Parent's voice conveys positive feeling to child
Parent spontaneously praises child's qualities twice during visit
Parent caresses, kisses, or cuddles child during visit

HOME Language and Cognitive Stimulation, 14 and 24 months:
At least 10 books are present and visible
Muscle activity toys or equipment available
Push or pull toy available
Stroller or walker, kiddie car, scooter, or tricycle available
Parent provides toys for child during visit (interviewer observation)
Learning equipment appropriate to age – cuddly toys or role-playing toys
Learning facilitators – child-sized table, chair, or booster chair
Complex eye-hand coordination toys (such as stacking or nesting toys, Legos, blocks)
Toys for literature and music (such as rattles, musical toys)
Parent talks to child while doing household work
Parent reads stories to child at least 2 times weekly
Child has 3 or more books of his/ her own

HOME Language and Cognitive Stimulation, 36 months:
Child has record player and at least five children's records
Child has access to at least 10 children's books
At least 10 books are visible in the apartment
Child is encouraged to learn shapes
Child is encouraged to learn the alphabet
Parent converses with child at least twice during visit
Parent answers child's questions or requests verbally
Parent usually responds verbally to child's speech
Child is encouraged to learn colors
Child is encouraged to learn numbers
Parent uses complex sentence structure and vocabulary

HOME Verbal/Social Skills Scale, 14 and 24 months:
Parent's speech is distinct and audible
Parent initiates verbal exchange with visitor
Parent converses freely and easily

Items in Bayley Language Factor score at 24 months.[a]
Child says eight different words (elicited or observed)
Child uses a two-word utterance (observed)
Child imitates a two-word sentence (elicited or observed)
Child identifies at least two objects in photo named by observer (rabbit, bell, cube, car, triangle)
Child uses pronouns (elicited or observed)
Child points to five objects in photo named by observer (dog, shoe, cup, house, clock, book, fish, star, leaf, car)
Child names three objects (ball, picture book, cup, spoon, pencil)
Child uses a three-word sentence (observed)
Child makes a contingent utterance (observed)
Child names five objects in photo (shoe, dog, cup, house, clock, book, fish, star, leaf, car)
Child poses question (observed)
Child produces multiple-word utterance in response to picture book

[a] Pan et al. (2004).

References

Bailey, G., Thomas, E., 1998. Some aspects of african american vernacular phonology. In: Rickford, J.R., Bailey, G., Baugh, J. (Eds.), African American English: Structure, History, and Use. Routledge, London.
Bayley, N., 1993. Bayley Scales of Infant Development, 2nd ed.: Manual. The Psychological Corporation, Harcourt Brace & Company, New York.
Becker, G., 1965. A theory of the allocation of time. The Economic Journal 299 (75), 493–517.
Brooks-Gunn, J., Klebanov, P., Duncan, G., 1996. Ethnic differences in children's intelligence test scores: role of economic deprivation, home environment, and maternal characteristics. Child Development 67 (2), 396–408.

Clopper, C.G., Pisoni, D.B., 2004. Some acoustic cues for the perceptual categorization of american english regional dialects. Journal of Phonetics 32, 111–140.
Duncan, G., Brooks-Gunn, J., Klebanov, P., 1994. Economic deprivation and early childhood development. Child Development 65 (2), 296–318.
Dunn, L., Dunn, L.M., 1997. Peabody Picture Vocabulary Test, Third ed. American Guidance Service, Circle Pines, MN.
Farkas, G., Beron, K., 2004. The detailed age trajectory of oral vocabulary knowledge: differences by class and race*. Social Science Research 33 (3), 464–497.
Feldman, H.M., Dollaghan, C.A., Campbell, T.F., Kurs-Lasky, M., Janosky, J.E., Paradise, J.L., 2000. Measurement properties of the macarthur communicative development inventories at ages one and two years. Child Development 71 (2), 310–322.
Fenson, L., Dale and J Reznick, P., Thal, D., Bates, E., Hartung, J.P., Reilly, J.S., 1993. The MacArthur Communicative Development Inventories: Users Guide and Technical Manual. Singular Publishing Group, San Diego, CA.
Fenson, L., Pethick, S., Renda, C., Cox, J.L., Dale, P.S., Reznick, J.S., 2000. Short-form versions of the macarthur communicative development inventories. Applied Psycholinguistics 21 (01), 95–116.
Fryer Jr., R., Levitt, S.D., 2004. Understanding the black–white test score gap in the first two years of school. Review of Economics and Statistics 86 (2), 447–464.
Guryan, J., Hurst, E., Kearney, M., 2008. Parental education and parental time with children. Journal of Economic Perspectives 2 (22), 23–46.
Hart, B., Risley, T.R., 1995. Meaningful Differences in the Everyday Experiences of Young American Children. Paul H. Brookes, Baltimore, MD.
Hoff, E., 2003. The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. Child Development 74 (5), 1368–1378.
Hoff, E., Naigles, L., 2002. How children use input to acquire a lexicon. Child Development 73 (2), 418–433.
Hoff-Ginsberg, E., 1991. Mother–child conversations in different social classes and communicative settings. Child Development 62 (4), 728–796.
Hoff-Ginsberg, E., 1992. How should frequency in input be measured? First Language 12 (36), 233.
Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., Lyons, T., 1991. Early vocabulary growth: relation to language input and gender. Developmental Psychology 27 (2), 236–248.
Huttenlocher, J., Vasilyeva, M., Waterfall, H.R., Vevea, J.L., Hedges, L.V., 2007. The varieties of speech to young children. Developmental Psychology 43 (5), 1062–1083.
Jann, B., 2008. A stata implementation of the blinder–oaxaca decomposition. The Stata Journal 8 (4).
Jones, F.L., Kelley, J., 1984. Decomposing differences between groups: a cautionary note on measuring discrimination. Sociological Methods & Research 12 (323).
Kalil, A., Ryan, R., Corey, M., forthcoming. Diverging destinies: maternal education and the developmental gradient in time with children. Demography.
Labov, W., 1972. Language in the Inner City: Studies in the Black English Vernacular. University of Pennsylvania Press, Philadelphia.
Love, J., Kisker, E., Ross, C., Schochet, P., Brooks-Gunn, J., Paulsell, D., Boller, K., Constantine, J., Vogel, C., Fuligni, A., 2001. Building their Futures: How Early Head Start Programs are Enhancing the Lives of Infants and Toddlers in Low-income Families, vol i. Technical report. eric.ed.gov.
Love, J., Kisker, E., Ross, C., Schochet, P., Brooks-Gunn, J., Paulsell, D., Boller, K., Constantine, J., Vogel, C., Fuligni, A., 2002. Making a Difference in the Lives of Infants and Toddlers and their Families: The impacts of Early Head Start, vols. i–iii: Final Technical Report [and] Appendixes [and] Local Contributions to Understanding the Programs and their Impacts.
Love, J.M., Kisker, E.E., Ross, C., Raikes, H., Constantine, J., Boller, K., Brooks-Gunn, J., Chazan-Cohen, R., Tarullo, L.B., Brady-Smith, C., Fuligni, A.S., Schochet, P.Z., Paulsell, D., Vogel, C., 2005. The effectiveness of early head start for 3-year-old children and their parents: lessons for policy and programs. Developmental Psychology 41 (6), 885–901.
Lugo-Gil, J., Tamis-LeMonda, C., 2008. Family resources and parenting quality: links to children's cognitive development across the first 3 years. Child Development 79 (4), 1–21.
Magnuson, K.A., Duncan, G.J., 2006. The role of family socioeconomic resources in the black–white test score gap among young children. Developmental Review 26 (4), 365–399.
Oaxaca, R., 1973. Sex discrimination in wages. In: Ashenfelter, O., Rees, A. (Eds.), Discrimination in Labor Markets. Princeton University Press, pp. 124–151.
Pan, B.A., Rowe, M., Spier, E., Tamis-Lemonda, C., 2004. Measuring productive vocabulary of toddlers in low-income families: concurrent and predictive validity of three sources of data. Journal of Child Language 31, 587–608.
Pan, B.A., Rowe, M.L., Singer, J.D., Snow, C.E., 2005. Maternal correlates of growth in toddler vocabulary production in low-income families. Child Development 76 (4), 763–782.
Pungello, E.P., Iruka, I.U., Dotterer, A.M., Mills-Koonce, R., Reznick, J.S., 2009. The effects of socioeconomic status, race, and parenting on language development in early childhood. Developmental Psychology 45 (2), 544–557.
Reznick, J.S., 1990. Visual preference as a test of infant word comprehension. Applied Psycholinguistics 11, 145–166.
Rickford, J.R., Rafal, C.T., 1996. Preterite had +v-ed in the narratives of african–american preadolescents. American Speech 71 (3), 227–254.
Rodriguez, E.T., Tamis-LeMonda, C.S., 2011. Trajectories of the home learning environment across the first 5 years: associations with children's vocabulary and literacy skills at prekindergarten. Child Development 82 (4), 1058–1075.
Rodriguez, E.T., Tamis-LeMonda, C.S., Spellmann, M.E., Pan, B.A., Raikes, H., Lugo-Gil, J., Luze, G., 2009. The formative role of home literacy experiences across the first three years of life in children from low-income families. Journal of Applied Developmental Psychology 30 (6), 677–694.
Rowe, M., 2008. Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. Journal of Child Language 35, 185–205.
Rowe, M.L., Pan, B.A., Ayoub, C., 2005. Predictors of variation in maternal talk to children: a longitudinal study of low-income families. Parenting 5 (3), 259–283.
Ware, A., Brady-Smith, C., O'Brien, C., Berlin, L., 1998. Child–parent Interaction Rating Scales for the Three-bag Assessment (14 months). Unpublished Scales. National Center for Children and Families, Teachers College, Columbia University.
Washington, J.A., Craig, H.K., 2002. Morphosyntactic Forms of African American English Used by Young Children and their Caregivers, vol. 23, pp. 209–231.
Weizman, Z., Snow, C., 2001. Lexical input as related to children's vocabulary acquisition: effects of sophisticated exposure and support for meaning. Developmental Psychology 37 (2), 265–279.
Winsborough, H.H., Dickenson, P., 1971. Components of negro-white income differences. Proceedings of the American Statistical Association, Social Science Section, 6–8.
Wolfram, W., 1971. Black–white speech differences revisited. In: Wolfram, W., Clarke, N.H. (Eds.), Black White Speech Relationships. Center for Applied Linguistics, Washington, DC.
Wolfram, W. 1991. Dialects and American English.
Woodcock, R.W., 1987. Development and Standardization of the Woodcock–Johnson Psycho-educational Battery. Teaching Resources, Hingham, MA.