


This article is commented on by Golinkoff et al (2019) and the authors in response. The commentary and author response were published in the previous 90.3 issue in error and are available online: Golinkoff et al (2019): <https://doi.org/10.1111/cdev.13128> Sperry et al (2019): <https://doi.org/10.1111/cdev.13125>

Reexamining the Verbal Environments of Children From Different Socioeconomic Backgrounds

Douglas E. Sperry 
Saint Mary-of-the-Woods College

Linda L. Sperry
Indiana State University

Peggy J. Miller
University of Illinois at Urbana–Champaign

Amid growing controversy about the oft-cited “30-million-word gap,” this investigation uses language data from five American communities across the socioeconomic spectrum to test, for the first time, Hart and Risley’s (1995) claim that poor children hear 30 million fewer words than their middle-class counterparts during the early years of life. The five studies combined ethnographic fieldwork with longitudinal home observations of 42 children (18–48 months) interacting with family members in everyday life contexts. Results do not support Hart and Risley’s claim, reveal substantial variation in vocabulary environments within each socioeconomic stratum, and suggest that definitions of verbal environments that exclude multiple caregivers and bystander talk disproportionately underestimate the number of words to which low-income children are exposed.

Recently, considerable attention has been paid to the disparity in the number of words spoken to very young children from socioeconomically disadvantaged families compared with their privileged peers. Although this relationship has been oft noted, the current reiterations of the argument herald the 30-million “Word Gap,” citing Hart and Risley’s (HR) study of 42 Kansas families conducted in the 1980s (Hart & Risley, 1995, 2003). HR calculated the mean number of words spoken to each child across 1-hr monthly observations from the child’s

first to third birthday. HR then extrapolated from this average to predict the number of words the child would hear in the first 4 years of life. In this manner, HR estimated that the children in their most impoverished group (six African Americans) heard 30 million fewer words than did the children in the most privileged group (13 offspring of professional families, one of whom was African American). These findings have never been replicated.

The Word Gap has garnered widespread attention beyond the academy. Media interest accelerated in 2013 when the Bloomberg Philanthropies Mayor’s Challenge awarded Providence, RI its grand prize to “Providence Talks.” This project proposed to teach poor parents how to speak to their children with the aid of an electronic device that measured their words. Another prominent example is the Clinton Foundation’s Too Small to Fail

This research was supported by a Spencer/National Academy of Education Dissertation Fellowship for Research Related to Education to Douglas E. Sperry and by a grant from the University of Illinois Research Board (Arnold O. Beckman Award) to Peggy J. Miller. We thank the families who generously participated in this research; the members of D.S.’s dissertation committee (Anne Haas Dyson, Cynthia L. Fisher, Wendy L. Haight, and Michèle Koven); and R. Bryant, M. Olivarez, C. Rundel, E. Siegel, E. E. Sperry, and A. Vowell for their assistance in transcription and coding.

Correspondence concerning this article should be addressed to Douglas E. Sperry, Department of Social and Behavioral Sciences, Saint Mary-of-the-Woods College, Saint Mary of the Woods, IN 47876. Electronic mail may be sent to dsperry@smwc.edu.

© 2018 Society for Research in Child Development
All rights reserved. 0009-3920/2019/9004-0020
DOI: 10.1111/cdev.13072

Initiative, which hosted the White House Word Gap Event in October 2014. Initiatives announced at this policy forum included potential funding by the U.S. Department of Health and Human Services for remedial efforts to address the Word Gap. Recent attention includes an Associated Press report on the Providence initiative that was picked up by major news outlets including the *New York Times* (Neergaard, 2017).

Not all attention to the Word Gap has been favorable. Some scholars suggest that the Word Gap is only the most recent vestige of a tendency to consider non-mainstream ways of speaking as deficient, using social class as a proxy variable for ethnic differences (Dudley-Marling & Lucas, 2009; Miller & Sperry, 2012). The validity of addressing only the number of words poor children hear in the preschool years as the cause of their inadequate school achievement has been questioned by anthropologists, linguists, and educators (e.g., Avineri & Johnson, 2015; Blum, 2015; Johnson, 2015). Critics charge that the Word Gap ignores the culturally defined contexts in which language is learned and used, and they take issue with the assumption that maternal vocabulary spoken directly to the child is the only speech that matters for language learning (Avineri & Johnson, 2015; Miller & Sperry, 2012; Zentella, 2015). They note that the practice of talking to the child in dyadic interaction is socioculturally defined (Brown & Gaskins, 2014; Duranti, Ochs, & Schieffelin, 2012; Miller, Cho, & Bracey, 2005), does not exist in many cultures (Brown & Gaskins, 2014; Duranti et al., 2012; Rogoff, 2003), and is not necessary for language learning (Akhtar & Gernsbacher, 2007).

Paralleling these conceptual criticisms are related methodological issues. Although the aim of HR was “to record ‘everything’ that went on in children’s homes” (Hart & Risley, 1995, p. 24), the methods HR employed may have eliminated adult talk that was not directed to the child. For example, they avoided recording family interactions not involving the child (p. 34) and tried not to encourage adult–adult conversation “for ease of transcription” (p. 39). HR also urged their observers to interact as little as possible with family members during the period of observation. These attempts at experimental control may have sacrificed ecological validity by making some participants more ill at ease than others (Dudley-Marling & Lucas, 2009). In addition, it appears that HR considered the language of only one parent, the child’s mother, in their estimates, although they sometimes referred generically to “parents.” Further, HR reported no

data about the number of words to which children were exposed in their ambient environments, that is, language addressed to other individuals but overheard by language-learning children. Although they stated that extended family members and other children were present at many observations (p. 31) and also reported the number of utterances addressed by other family members to the child and by other family members to each other (Hart & Risley, 1999), HR did not include the language of these parties in their estimates of the Word Gap. Their data reduction process involved “picking out the child and parent from all the other conversations going on at the same time” (Hart & Risley, 1995, p. 41). Despite the fact that HR reported that nearly half of the utterances their child participants heard were not addressed to them (Hart & Risley, 1999, p. 34), HR’s methodological decisions effectively defined and operationalized the vocabulary environment as speech directed to the child by the primary caregiver, excluding speech by other family members to the child as well as speech that the child overheard.

In other words, the gap in the number of words addressed to children from different social classes reported by HR reflects only a subset of the total number of words in the Kansas children’s verbal environments. This fact raises critically important questions, particularly given recent evidence in the psycholinguistic and language socialization literatures that interrogates the relative contributions and benefits of overheard versus directed speech. At the most fundamental level, little is known about the full verbal environment in the homes of children from different social class backgrounds. The study presented here addresses this question.

In the face of this controversy, it may be helpful to step back and assess what is known about vocabulary learning in the preschool years. There is substantial evidence that children who hear more language learn more language (Hoff, 2003; Hurtado, Marchman, & Fernald, 2008; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). The amount of language children hear may be measured in terms of the number of unique words (word types) or the total number of words (word tokens) spoken by caregivers. Although much recent scholarship has focused on word types (e.g., Hoff, 2003; Huttenlocher, Vasilyeva, Waterfall, Vevea, & Hedges, 2007; Pan, Rowe, Singer, & Snow, 2005; Rowe, 2008), syntactic structures (Huttenlocher et al., 2007), and discourse features (Hoff, 2003), the Word Gap was measured in terms of the total number of word *tokens* addressed to children. This situation represents a curious

anomaly because the initial work of HR addressed quality measures such as the interactional features of utterances (e.g., whether utterances were repetitions, paraphrases, or expansions), and the degree to which such utterances questioned the child's knowledge or prohibited the child's actions (Hart & Risley, 1992, 1995, 1999).

Joint attention between a caregiver and a child has often been assumed necessary for word learning. Tomasello (1995) defined joint attention to include only those interactions where children alternated their gaze between the adult stating an object label to be learned and the entity being labeled. Recent evidence has challenged this view. For example, Akhtar and Gernsbacher (2007) suggested that one of the key problems with the joint attention model of word learning is that it relies too heavily on the process of overt attention. Akhtar (2005) examined the robustness of vocabulary learning through overhearing by testing 2 year olds in contexts where a potentially distracting activity was present. Children watched while the experimenter and a confederate examined four novel objects, only one of which they named ("toma" or "modi"). Children either played with another interesting toy (the distracter condition) or simply observed the experimenter's interaction with the confederate (no-distracter condition). At test, all children were significantly more likely to choose the target, labeled object in both conditions despite the fact that children in the distracter condition had paid significantly less attention to the experimenter than children in the no-distracter condition.

Shneidman, Sootsman Buresh, Shimpi, Knight-Schwarz, and Woodward (2009) examined whether children would learn novel words in the absence of joint attention between infant and speaker, and also asked whether or not this ability is correlated with the amount of ambient speech their participants heard every day in their homes. They engaged 20 month olds in one of two tasks: a direct condition where infants were shown an object by the experimenter while she labeled it ("Look at the blicket"), and the overhearing condition where infants watched while one experimenter both showed and labeled an object to another experimenter without making eye contact with the infant. In subsequent testing, children in both conditions succeeded at selecting the labeled object. Furthermore, successful word learning was positively correlated with the amount of time spent looking at the experimenters and negatively correlated with the amount of time spent looking at the object, suggesting that children may search for behavioral cues

in the actions of others that might signal the focus of their conversations.

Children can also encode syntactic-structure information and use it to interpret novel verbs in the absence of ostensive cues to verb meaning in a manner consistent with the type of learning that would be fostered by overheard speech (Arunachalam, Escovar, Hansen, & Waxman, 2013; Arunachalam & Waxman, 2010; Messenger, Yuan, & Fisher, 2015; Scott & Fisher, 2009; Yuan & Fisher, 2009). If children learn word meanings from overheard speech, they must be able to extract cues from their environment that signal those meanings in the absence of the sort of direct reference available through joint attention. One method by which children accomplish this feat may be through the use of distributional cues to infer the meaning of novel verbs (Scott & Fisher, 2009). Furthermore, 21-month-old children represent transitive properties versus intransitive properties of a novel verb based only on syntactic information presented in non-ostensive dialogs even before they have mastered transitive structures in their own speech (Arunachalam et al., 2013). Messenger et al. (2015) demonstrated that such learning may persist in long-term memory.

Despite strong experimental evidence that overheard speech is sufficient for language learning of various stripes, it remains that episodes of directed speech (or joint attention) make significant contributions to early word learning, even in communities in which overheard speech predominates. In a cross-cultural study, Shneidman and Goldin-Meadow (2012) compared the relative amounts of directed versus overheard speech in the everyday conversations of Yucatec Mayan and American children. Both groups of children and their families were followed longitudinally throughout their second and third years. At all data points, Mayan children heard more overheard than directed speech, although the amount of directed speech they heard grew steadily across this time frame. Nevertheless, only the number of word types in directed input at 24 months predicted vocabulary knowledge at 35 months. Neither the number of word types in overheard input nor the combination of word types in overheard and directed input accounted for a significant amount of variance in the number of word types in child speech.

The relative overall paucity of speech in the Mayan community as compared to the American community, however, calls into question the relative importance of directed versus overheard speech for children who hear a far greater amount of

input. Shneidman, Arroyo, Levine, and Goldin-Meadow (2013) addressed this issue in a study of 30 Chicago children and their families, ages 14–42 months, in longitudinal observations of naturally occurring familial interaction. The families were chosen from a larger investigation of language development based on the categorization of child language input as stemming from majority multi-speaker versus single-speaker interactions. Children in both types of homes heard the same amount of directed speech addressed to them by their primary caregiver; in addition, there was no significant difference in the amount of child-directed speech by primary caregivers in single-speaker households compared to child-directed speech by all interlocutors in multispeaker households. However, children in multispeaker homes did hear a greater number of both word types and tokens when overheard speech was added to the mix. Only child-directed speech (and not overheard speech) at 30 months predicted vocabulary development at 42 months for children from either household type.

Fernald and her colleagues (Fernald, Marchman, & Weisleder, 2013; Fernald, Perfors, & Marchman, 2006; Hurtado et al., 2008; Weisleder & Fernald, 2013) have provided evidence for the potential mechanism behind the relationship between vocabulary directed to the child and later academic achievement, namely that increased amounts of vocabulary addressed to the child are associated with increased speech processing ability. Effects are bidirectional, with increased amounts of language addressed to the child correlating with speech processing ability as young as 18 months (Hurtado et al., 2008) and with increased speech processing predicting vocabulary and grammatical development in the second year of life (Fernald et al., 2006) and language outcomes in elementary school (Marchman & Fernald, 2008). Importantly, in a study of 29 Latino children ages 19 and 24 months, Weisleder and Fernald (2013) demonstrated that the amount of directed, but not overheard, speech these children heard at 19 months predicted both language processing efficiency and vocabulary size at 24 months. Furthermore, children's language processing efficiency at 19 months was demonstrated to mediate the relationship between child-directed speech at 19 months and vocabulary outcomes at 24 months.

Complicating this picture further is the finding that various aspects of speech directed to the child change in importance across the early language learning years. Rowe (2012) found that when children were 18 months old, the quantity of parental

speech (word tokens) directed to the child contributed to vocabulary achievement at 30 months but declined in importance as children grew older. Vocabulary diversity (word types) heard by children at 30 months predicted vocabulary achievement at 42 months, and the number of decontextualized utterances heard by children at 42 months predicted vocabulary achievement at 54 months. These findings allude to the possibility that the respective roles of directed versus overheard speech may change as well.

In sum, the foregoing literature from the psycholinguistic tradition has produced strong evidence that young children can learn vocabulary from directed speech *and* from overheard speech. It also shows that speech directed to children predicts later vocabulary growth and language outcomes in school and suggests that different features of directed speech may have different effects as children get older. Although several studies have examined relationships between overheard versus directed speech and children's later vocabulary knowledge, and have shown a correlation with directed speech only (Shneidman & Goldin-Meadow, 2012; Shneidman et al., 2013; Weisleder & Fernald, 2013), this literature tells us very little about overheard speech—its frequency in different sociocultural groups, its features, or its possible benefits. These issues are not trivial, particularly because one of the most robust findings from cross-cultural research is the prevalence and efficacy of observational learning across a host of developmental domains (e.g., Gaskins & Paradise, 2010; Lancy, 2015; Miller & Cho, 2018; Ochs & Schieffelin, 1984; Rogoff, Paradise, Arauz, Correa-Chávez, & Angelillo, 2003). In short, the American middle-class model is anomalous within the global context (Henrich, Heine, & Norenzayan, 2010; Miller & Cho, 2018). Most relevant to the issues at hand, the interdisciplinary field of language socialization has investigated the cultural organization of childrearing, verbal interaction, and language learning in diverse cultures and communities within the United States and around the world since its inception in the 1980s. The evidence from this body of work runs parallel to the foregoing summary of cross-cultural research: The middle-class model of dyadic speech in which the parent speaks directly to the child and treats him as a conversational partner is neither common nor necessary for language learning (Duranti et al., 2012; Ochs & Schieffelin, 1984). In societies in which young children are seldom spoken to, they nonetheless reach major milestones of language development at ages comparable to those of Western

children (Brown & Gaskins, 2014). Moreover, these sources attest that most young children spend their daily lives in multiparty social constellations that afford a range of normative participant roles (Goffman, 1981), especially bystander and overhearer roles.

These roles seem especially congenial to certain genres of talk, notably oral narrative. In a study of personal storytelling, Miller, Fung, Lin, Chen, and Boldt (2012) found that young children from middle-class Taiwanese families in Taipei and middle-class European American families in Chicago had routine access to both co-narrator and bystander roles. However, the Taipei children not only occupied the bystander role much more frequently than the Chicago children but also engaged in twice as much listening from the bystander perspective across the entire age range (2;6–4;0). These patterns suggest that the Taiwanese families did not treat child listening as an immature response to be outgrown but as a valued activity, worthy of cultivation, findings that dovetail with other studies (cf. Li, 2012). Personal storytelling is even more common in poor and working-class communities in the United States. In these homes, this genre is highly valued and avidly practiced, the verbal environment is densely populated with stories, and young children participate by observing, listening, and co-narrating, developing precocious narrative skills (Miller et al., 2005). Stories go on around, about, and to young children (Heath, 1983; Miller & Sperry, 2012; Ward, 1971), and children hone their “nosiness” by listening in to their elders (Hudley, Haight, & Miller, 2003; Morgan, 1980). In one study of working-class African American families, personal stories accounted for one-quarter of 2 year olds’ naturally occurring speech (Sperry & Sperry, 1996, 2000), a wealth of decontextualized talk that may bode well for their vocabulary development (Rowe, 2012).

The availability in poor and working-class families of language, in general, and of structurally and representationally complex forms of language such as narrative, in particular, cause us to echo the concerns of others (Callanan & Waxman, 2013; Cole, 2013) about the link between amount of speech in young children’s verbal environments and social class. Although social class has been used as a proxy variable for more intricate processes and interactions between parent and child, it does not appear to be, in and of itself, an environmental influence on later development. Hart and Risley (1999) and Hoff (2003) demonstrated that when relations exist between social class and early vocabulary development, they

are mediated by particular properties of maternal speech (such as number of utterances, word tokens, word types, and topic-continuing replies) and that these properties and their effects operate with efficient specificity. Furthermore, not all studies have shown that maternal talkativeness correlates with child vocabulary growth (Pan et al., 2005). Many studies have reported a wide range of maternal talkativeness within low-income families (Hurtado et al., 2008; Rowe, Pan, & Ayoub, 2005; Weisleder & Fernald, 2013), suggesting that verbal experience is not uniform within any social class. In fact, maternal education has been shown to predict language growth independently of socioeconomic status (Hoff, 2006, 2013; Huttenlocher et al., 2007; Rowe, 2008, 2012). What remains unclear is the degree to which the many properties of maternal language conducive to later language learning are located exclusively in the speech of any group of mothers defined by social class or ethnicity, particularly given the diversity of beliefs concerning childrearing in general, and language and communication practices in particular, that exist across the range of American communities (e.g., Kusserow, 2004; Miller et al., 2005).

Having considered relevant studies from the psycholinguistic and language socialization literatures, we now return to our starting point: the controversy surrounding HR’s finding of a Word Gap between the number of words spoken to young children in poor families, compared with privileged families. The foregoing literatures underwrite the need to take another look at HR’s claim and to do so with an eye to variation in children’s verbal environments, especially with respect to overheard speech versus speech directed to the child and to dyadic versus multiparty configurations. Given these concerns, the study reported here addresses whether or not the social class differences observed by HR obtain across a different sample of American communities. It also addresses the number of words heard by children under three different definitions of the verbal environment: (a) HR’s definition (speech by the primary caregiver to the child), an expanded definition that includes (b) all speech directed to the child, and a further expanded definition that includes (c) bystander speech in addition to speech directed to the child.

Method

This study examines extant corpora of language data from investigations conducted in five American communities from five time periods extending

Table 1
Description of Communities, Participants, and Transcribed Data Corpora

Site	Participant SES and ethnicity	Extended family contact	Number and gender of participants	Age range of observations (in months)	Total number of transcribed samples	Length of transcribed samples (in minutes)	Total transcribed data (in hours)
South Baltimore	Poor European American	Extensive	3 girls	18–32	35	60	35
Black Belt of Alabama	Poor African American	Extensive	5 boys 6 girls	24–42	64	30	32
Jefferson, Indiana	Working-class European American	Moderate	8 boys 7 girls	18–42	135	30	67.5
Daly Park (Chicago)	Working-class European American	Moderate	4 boys 3 girls	30–48	26	30	13
Longwood (Chicago)	Middle-class European American	Limited	3 boys 3 girls	30–48	20	30	10
Total			20 boys 22 girls	18–48	280	30–60	157.5

Note. SES = socioeconomic status.

from the late 1970s through the late 1990s (see Table 1). These corpora were selected on the basis of two criteria. First, the original studies were designed to document everyday talk and interaction in families with young children and to do so in a manner that was ecologically and culturally valid. By employing both ethnographic fieldwork and word counts based on longitudinal home observations, this study follows the growing trend in developmental psychology of combining quantitative and qualitative methods to study development in context (e.g., Duncan, Huston, & Weisner, 2008; García Coll & Marks, 2009; Rosengren et al., 2014; Sperry & Sperry, 2016; Weisner, 2005). Second, in parallel with HR, the communities in this study represent a spectrum of socioeconomic statuses, with a total sample size ($N = 42$) identical to that of HR. The spectrum includes 14 poor children from two communities (HR had 6 welfare children), 22 working-class children from two communities (HR had 13 lower class children), and 6 middle-class children from one community (HR had 10 middle-class children). HR also had 13 children from “professional” families; we had none. Thus, our sample has more poor and working-class children and fewer middle- and upper class children compared with HR.

Our five studies adopted the same methodological approach: Extensive ethnographic fieldwork,

followed by longitudinal home observations (more methodological detail is provided in Miller, 1982; Sperry & Sperry, 1996; Wiley, Rose, Burger, & Miller, 1998). In the fieldwork phase, which lasted at least a year, researchers spent a considerable amount of time getting to know the community, developing a network of contacts, and visiting institutions (e.g., health clinics, preschools) that had positive relationships with potential participants. In keeping with standard ethnographic practice (e.g., Jessor, Colby, & Shweder, 1996; Miller, Hengst, & Wang, 2003; Wolcott, 1995), the ethnographer tried to fit in with local ways, navigating differences of social class and ethnicity, and negotiating a role that was comfortable and culturally appropriate. For example, researchers participated, upon request, in other community contexts (e.g., tutoring children on school work). Once the families were recruited, researchers visited repeatedly to get acquainted and learn about daily routines before the observation phase began. As the study progressed, the families sometimes invited the researchers to family events or asked for assistance with transportation.

Participants

Families were recruited through contacts in the community and selected according to three criteria:

(a) the families were representative of their community and possessed concurrent and intergenerational ties to the community, (b) the children were the appropriate age, and (c) the children were developing normally. The children were 18–30 months of age when recordings began.

Communities

Tabular summation of demographic and sampling characteristics is presented in Table 1.

South Baltimore (Poor, European American, n = 3)

The three girls lived with their mothers, who received public assistance. Two of these children lived in extended family arrangements, whereas the third enjoyed frequent visits from other family members.

The Black Belt (Poor, African American, n = 11)

Nine families qualified for free- or reduced-lunch and food stamps, and five received public assistance. Five families lived in public housing. All but one of the children lived in extended family arrangements, with family members residing under the same roof or in nearby homes which children visited frequently. Five families had at least one parent holding an unskilled job, and one mother was a public school teacher. Six families did not have an adult with gainful employment at the time of the study.

Jefferson (Working Class, European American, n = 15)

One or both parents held unskilled jobs. All families qualified for free- or reduced-price lunch. All the children lived in nuclear family arrangements and had routine contact with friends and relatives.

Daly Park (Working Class, European American, n = 7)

The fathers were employed in blue-collar jobs; only one mother worked outside the home. Three of the children had at least one parent who had attended some college or had received a college degree. All the children lived in nuclear family households. Only one family owned their own home.

Longwood (Middle Class, European American, n = 6)

All fathers but one were employed in white-collar jobs; the remaining father had left his white-

collar job to be a police officer to avoid being transferred overseas. The mothers self-identified as “stay-at-home moms.” All the parents had college degrees, and each family owned its own home.

Procedures

The observational phase of each study consisted of a series of regular longitudinal home observations. The frequency and length of observations varied from study to study but were consistent within each study. Children were videotaped interacting with all interlocutors present. The principal caregiver (mother, grandmother, or father) was present; often the child participant’s siblings or other members of her extended family (e.g., cousins, teen-aged aunts and uncles) were also present. For the most part, observations were conducted in a central room, usually the living room.

Observation sessions were characterized by the frequent comings and goings of adults and children, conversations about school and work days, and everyday speech surrounding quotidian acts such as meal preparation and homework. The researchers did not make any attempt to alter the family’s behavior or dictate their activities nor did they strive to be invisible. They tried to interact with adults and children in an interested and relaxed manner, much as a family friend might do. To the greatest extent possible, these data represent snapshots of American families doing what they do, every day, in their own homes.

The observations were transcribed verbatim. In South Baltimore complete transcripts of the 1-hr observations were available. In the other cases, the second half-hour of each observation was transcribed. This selection was made to avoid times of unusual excitement or fatigue on the part of the child. Transcripts contained detailed nonverbal and contextual cues, consistent with the goal of examining the total ambient environment of the language-learning child. Transcripts were checked at least twice. In the case of the Black Belt, the regional dialectal variation of African American Vernacular English required special attention; a college student from the community made the initial transcript.

Coding

For the present analysis, a total of 280 observations, comprising 157.5 hr of family interaction across 42 children, were examined (see Table 1). Language data were first sorted by speaker and addressee. Speaker categories included child

participant, primary caregiver (usually the mother), youth, other adult, and researcher. Addressee categories included child participant, other, and researcher. When two or more children were present, adults sometimes jointly addressed both children. In those cases, the speech was counted as addressed to the focal child. Words were not counted in multiple categories. Finally, the visits of the researcher, like those of other guests, often engendered considerable talk (e.g., “catching up” since the last visit) by family members and focal children. To ensure a conservative estimate of the verbal environment, speech by the researcher and addressed specifically to the researcher was discarded from further analysis.

The number of word tokens in each speaker—addressee file was then calculated using CLAN (Computerized Language Analysis) within CHILDES (Child Language Data Exchange System; MacWhinney, 2000). For the present analysis we used word tokens instead of word types to allow comparison to the metric used by HR to calculate the Word Gap. Type analyses for these data are available, however (Sperry, 2014). Decision rules for what constitutes a word or lexeme were developed in close consultation with published rules for counting vocabulary in mother–child talk (e.g., Huttenlocher et al., 1991; Rowe, 2008). Regular inflectional morphemes (regular verb endings, regular plural or diminutive noun endings) were reduced to their base lexeme (e.g., “walk,” “walks,” and “walked” were reduced to the lexeme WALK; “dog,” “dogs,” and “doggie” were reduced to the lexeme DOG). Verb catenatives (e.g., GOTTA) and other words involving clitics (e.g., HE’LL) were counted as unique words, independent of any instance of their component parts. Reduplicatives (e.g., BYE-BYE) and onomatopoeic sounds were reduced to one word. All dialectal phonetic equivalents were reduced to a standard spelling across transcripts from the different communities in order to allow for computer analysis. Final reductions in vocabulary were hand checked using CLAN output by the first author.

Results

Three analyses of the children’s verbal environment were undertaken, corresponding to the three definitions described earlier. In the case of the first definition, *words spoken by the primary caregiver to the focal child*, the results of this study are juxtaposed against the results of HR’s investigation in four Kansas

communities (Hart & Risley, 1995). The data for the Kansas communities consist of the original mean numbers of tokens reported by HR (1995, Appendix A, pp. 228–229). It should be noted that HR sometimes described their communities in terms of four social class groups and sometimes in terms of three groups, which they created by combining their “middle”-class and “lower” class families into a larger “working”-class group (1995, p. 32; 2003). In our analyses, we opted to use their four-way classification (1995, p. 31). This decision allowed us to “match” our groups to theirs more precisely on the basis of employment and educational status. Our middle-class group (Longwood) corresponded to their “middle-class” group, our working-class groups (Jefferson, Daly Park) to their “lower class” group, and our poor groups (South Baltimore, Black Belt) to their “welfare” group.

Analysis of Definition 1

The comparison between the mean number of words spoken *by the primary caregiver to the focal child* in each of our five communities and the number of words spoken to children in HR’s four Kansas communities (Definition 1) revealed no clear pattern in terms of social class (Figure 1). The means of the nine communities were compared using the Tukey–Kramer test of Paired Comparisons. Only the Kansas Professional to Kansas Welfare comparison reached statistical significance, $Q(9, 75) = 5.39, p < .01$. In other words, although there is reason to assume that the primary caregivers in the Kansas Professional community spoke more words to their children than did the primary caregivers in the Kansas Welfare community, there is no reason to assume that there are differences between any of the other communities, either between themselves or between them and the Kansas Professional and Kansas Welfare communities. Interestingly, the only other comparison to approach significance was that of the Black Belt and Kansas Welfare communities, $Q(9, 75) = 4.29, p < .08$. This fact, coupled with further examination of the descriptive data, suggests a very different pattern than that offered by HR. Although the Kansas Professional community and the Kansas Welfare communities remained at the extreme high and low of this distribution, the Black Belt children living in poor homes heard more words spoken by their primary caregivers to them than did children in any other community except the Kansas Professional community (including notably the middle-class communities of Longwood and Kansas). Although

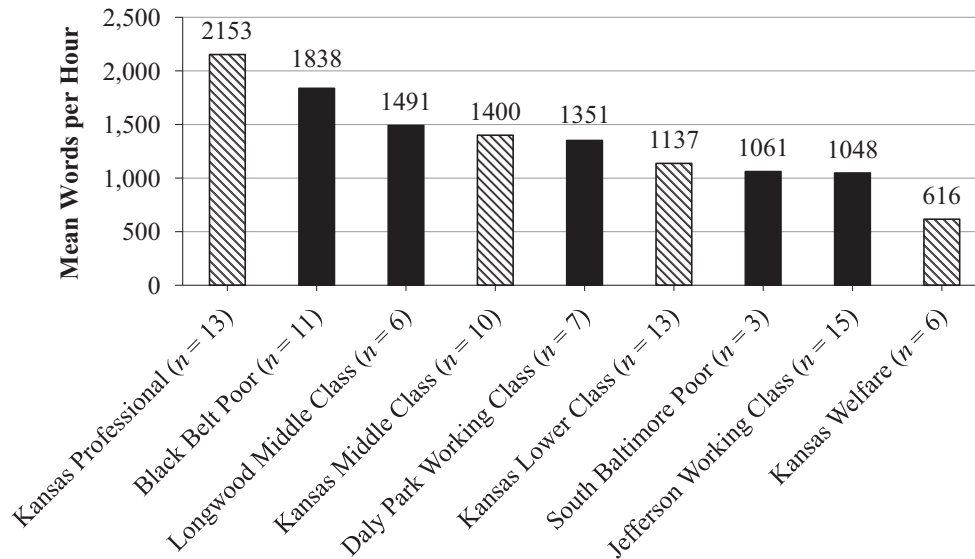


Figure 1. Mean number of words spoken by primary caregivers to children. Data collected for the present analysis are presented in solid bars. Data collected by Hart and Risley (1995) and used here for comparison are presented in hashed bars.

considerable similarity occurred in the number of words spoken by primary caregivers to their children in the working-class community of Daly Park and the Kansas lower class community, the number of words spoken by primary caregivers to their children in the working-class community of Jefferson was more equivalent to that of the poor community of South Baltimore. In sum, when HR's definition of the vocabulary environment was applied to the five communities in our study, we found that caregiver talkativeness was subject to community variation not predicted by socioeconomic status. Moreover, the finding of statistical difference only between the Kansas Professional and Kansas Welfare communities suggests that these two Kansas communities may be outliers.

Analysis of Definition 2

Our second analysis corresponded to Definition 2, *words spoken by all caregivers to the focal child*. As can be seen in Figure 2, the contrast between the number of words spoken to the child by the primary caregiver (Definition 1) versus the number of words spoken to the child by all caregivers (Definition 2) helped to demonstrate the perils of making assumptions concerning vocabulary quantity based on socioeconomic differences alone. The community-level comparisons shown in Figure 2 were additionally interrogated by examining the mean increase between Definitions 1 and 2 across individual families within each community (see Table 2;

note that mean percentage differences shown in Table 2 do not correspond to community-wide means in Figure 2 because they are based on average percentage increases, and not average total tokens, across individual families and conditions). These two analyses reveal that whereas primary caregivers in the middle-class community of Longwood spoke an average of 1,491 words per hour to the focal children, all caregivers in this community spoke an average of 1,777 words to the focal children, an increase of 30%. By contrast, the analogous increases in the working-class communities were 17% (Daly Park) and 53% (Jefferson); in the poor communities, they were 21% (South Baltimore) and 58% (Black Belt). One possible reason for these differences is the greater number of older siblings in some communities (e.g., Longwood, Jefferson, Black Belt) than in others (Daly Park, South Baltimore). These findings demonstrate that social class alone did not determine either the composition of households or, consequently, the amount of speech children heard.

Analysis of Definition 3

In our analysis of Definition 3, *all ambient speech within the child's hearing*, Figure 2 demonstrates that children in every community were exposed to far more language than that addressed specifically to them. In the poor community of South Baltimore, where extended family members were often present, the mean number of words the children heard

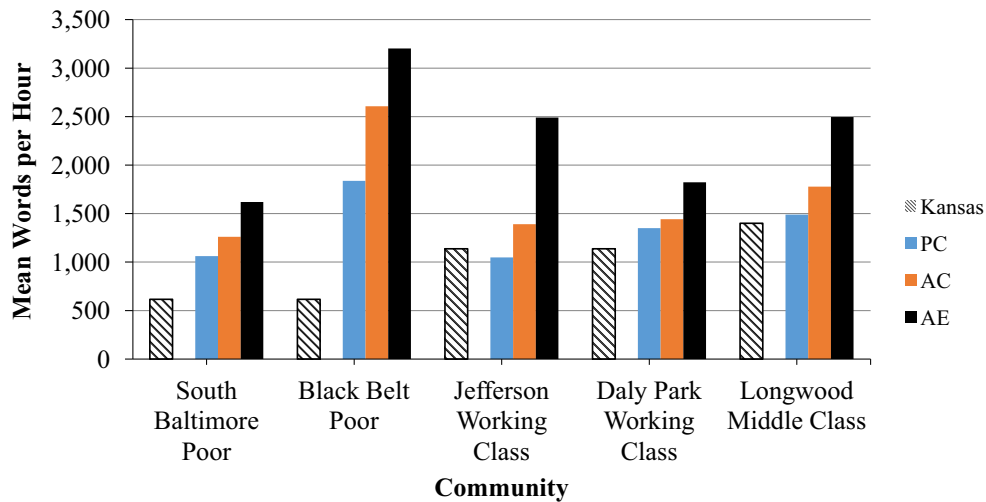


Figure 2. Mean number of words spoken across three definitions of the verbal environment: By the primary caregiver to child (PC); by all caregivers to child (AC); and all speech in child’s ambient environment (AE) to and around child. Data (speech by primary caregiver to child) from Hart and Risley (HR, 1995) included for comparison in hashed bars; note that HR did not report the number of words spoken by other interlocutors in the child’s environment. The HR data compared to the South Baltimore and Black Belt samples are those from the Kansas Welfare community. The HR data compared to the Jefferson and Daly Park samples are those from the Kansas Lower Class community. The HR data compared to the Longwood sample are those from the Kansas Middle Class community.

Table 2
Average Percentage Increases When Three Definitions of the Verbal Environment Are Compared: Primary Caregiver Speech to Child (Definition 1), All Caregiver Speech to Child (Definition 2), and All Speech in the Ambient Environment to and Around the Child (Definition 3)

	Definitions 1-2	Definitions 1-3	Definitions 2-3
South Baltimore Poor	21	54	27
Black Belt Poor	58	102	23
Jefferson Working Class	53	210	90
Daly Park Working Class	17	75	46
Longwood Middle Class	30	104	54

per hour represented a 54% increase over the number of words spoken to them by their primary caregivers alone (see Table 2). In the working-class community of Jefferson, the amount of ambient speech heard by children was a 210% increase over the number of words addressed to them by a single caregiver due to the presence of siblings and of both parents. Interestingly, the same amount of ambient speech was present in both the working-class community of Jefferson and the middle-class community of Longwood despite a substantial

difference between these two communities in the amount of speech addressed specifically to the child. Finally, children in the poor Black Belt community heard an astounding 3,203 words per hour in their ambient environment, a 102% increase over the number of words addressed to them by their primary caregivers. This mean exceeded both the group mean (2,153) of words spoken by primary caregivers to the child for the Kansas Professional families and the individual family means for 12 of the 13 Professional families.

Summary Analysis of Three Definitions of Verbal Environments

Figure 2 presents a summary view of our three definitions of the verbal environment along with a comparison to means of child-directed speech reported by HR. Although it was not possible to compare the results of our second and third analyses directly to HR because HR did not provide word token data from children’s total verbal environments, our results confirmed that the children in our five communities heard far more speech than HR’s results implied. For example, the Black Belt children living in poverty heard more words spoken to them per hour than any other children, fully 21% more words overall in their everyday interactions with family members than was reported for the children in the Kansas Professional homes. The children in working-

class Daly Park heard more words addressed to them than was reported for all the children in the Kansas samples except for the children of professional parents; the children in working-class Jefferson heard nearly as many words as were reported for the children in the middle-class Kansas homes.

Discussion

The study reported here represents the first attempt to replicate Hart and Risley's claim of a massive Word Gap in the vocabulary environments of poor children, compared to their more privileged peers. Our findings do not support their claim. When HR's definition of the vocabulary environment (speech of the primary caregiver to the child) was adopted, evidence for a relationship between social class and the number of words addressed to young children was weak. When more expansive definitions of the verbal environment were employed, definitions that derive from the scholarship reviewed earlier, the evidence pointed in a different direction. Not only did the Word Gap disappear, but also some poor and working-class communities showed an advantage in the number of words children heard, compared with middle-class communities. Our study also revealed a great deal of variation among communities *within* each socioeconomic stratum, consistent with earlier studies (Hurtado et al., 2008; Rowe et al., 2005).

How can we explain the discrepancy between HR's findings and our own? Our failure to replicate HR's findings when using their definition of the vocabulary environment raises the possibility that variation across communities within a particular social class is so great that it swamps variation across classes. This possibility gains weight when sample sizes per community are small, as they were in both studies, ranging from 3 to 15 in our study and from 6 to 13 in HR's study. It is also the case that the socioeconomic spectrum is not identical in the two studies: Our study is more heavily weighted toward the lower end of the spectrum, with two poor communities and two working-class communities but only one middle-class community and no professional community. By contrast, HR included both a middle-class and a professional sample but only one poor and one lower class sample. However, the fact that the present sample is more heavily weighted with lower income participants supports the possibility that these data provide a more representative picture of verbal environments in these households, a significant

issue because our lower income families generally spoke greater numbers of words to their young children than did those families in HR's study.

There are also some differences between the two studies in the children's ages. Observations began at 18 months in South Baltimore and Jefferson, at 24 months in the Black Belt, and at 30 months in Daly Park and Longwood; in all but the South Baltimore corpus the original data collection was geared to the study of oral narrative and thus began later and did not sample as frequently. This inconsistency of sampling when compared with other quasi-experimental studies focused on vocabulary input is a common issue in corpora studies (Goodman, Dale, & Li, 2008). The fact that HR's data collection began at 12 months opens the possibility that our data overestimate the number of words heard by children because our participants are slightly older, although recent reports suggest that the amount of speech addressed to children remains relatively constant across this age range (Huttenlocher et al., 2007; Shneidman & Goldin-Meadow, 2012). By contrast, our later sampling provides perhaps a better glimpse at the role of overheard speech in the verbal environment, as the role of directed speech diminishes as children become more competent language users (cf. Hart & Risley, 1999).

In further consideration of the discrepancy between the results of the two studies, there is another possibility, namely differences in methods. In some ways, our methods and HR's methods are similar: Both studies had modest total sample sizes of 42, and both collected enormous longitudinal samples of observed speech in children's homes, followed by detailed transcription. (It is hard to overemphasize just how labor intensive these methods are, which helps to explain why such studies are so rare.) However, unlike HR, we did not try to (a) avoid recording family interactions in which the child was present but did not participate verbally (e.g., adult-adult conversation) or (b) alter the natural interaction between observer and participants. These practices would have been at odds with the ethnographic approach that prioritizes local ecologies, cultural meaning systems, and normative language practices. As a result, it seems likely that our observations captured features of family life (e.g., multiple caregivers) and associated language environments (e.g., multiparty and bystander talk) that were precluded by HR, features that are common in low-income communities (Miller et al., 2005). In the end, regardless of methodological differences, we know little about vocabulary in the ambient

environments of the Kansas samples because, quite simply, HR did not report it.

In their review of the literature on language acquisition and language socialization, Brown and Gaskins (2014) addressed key issues related to what counts as relevant speech for vocabulary learning, noting the conflict between, on the one hand, findings that speech addressed to the child predicts lexical development and, on the other hand, findings that young children in societies where they are seldom spoken to nonetheless attain linguistic milestones at comparable rates. They suggested a resolution to this puzzle by reference to studies of societies in which children are socialized to attend intently to what is going on around them; such children may “be attuned to attend to others’ language and interactions, and be able to profit from overheard speech in ways unlike those of infants in societies where child-centered face-to-face interactions are the norm” (p. 201). Shneidman and Woodward (2016) made a similar argument: “Rather than signaling that child-directed interactions have universal, a priori information value, the empirical record suggests that children learn to see directed interactions as informative in some contexts based on their social experiences” (p. 13). The underlying principle in both arguments is that normative variation (in this case, child-directed vs. overheard speech) in the verbal environments of young children from different social and cultural addresses will likely instill different preferred normative strategies for attending to and learning from speech, which will in turn afford different benefits. This idea dovetails with growing evidence that different cultural practices yield different developmental trajectories and outcomes across a host of domains (e.g., Göncü, 1999; Goodnow, Miller, & Kessel, 1995; Miller et al., 2012; Ochs & Schieffelin, 1984; Rogoff et al., 2003; Shweder et al., 2006; Weisner, 2005).

Taken together, the foregoing considerations suggest that future research will need to address three analytically distinct areas in order to advance our understanding of sociocultural variation in early verbal environments and vocabulary learning. The first has to do with refining descriptions of the verbal environments that are routinely available to young children from different social class backgrounds, including identifying the nature and amount of speech in the ambient environment that might reasonably constitute overheard speech for language-learning children. This task is where the current paper makes its chief contribution by demonstrating that there is much more to the

verbal environment than speech directed to the child by a single caregiver, by contributing significant new findings to the growing body of evidence that multiparty and bystander configurations are widespread and preferred in many sociocultural communities, and by highlighting the variation *within* poor and working-class communities. With respect to variation in verbal environments, our findings point neither to a positive correlation between social class status and quantity of speech addressed to children (Definition 1) nor to any consistent relationship between social class and overheard speech (Definition 3). For example, one of our most intriguing findings is that primary caregivers from the Black Belt, a poor African American community, not only addressed far more words *to* children than did primary caregivers from other poor communities (our South Baltimore families and HR’s welfare families) but also exceeded the number of words produced by primary caregivers from every other higher status community in our study and HR’s study, with the exception of HR’s professional community. These findings suggest that future research should explore the sources of variation that differentiate communities *within and across* socioeconomic strata.

The second area has to do with determining the different kinds of attentional and learning strategies and associated benefits that are cultivated when children are routinely exposed to different kinds of social configurations and normative language practices other than predominantly child-directed speech (Brown & Gaskins, 2014). For example, when referring to recent studies that show relations between child-directed speech and later vocabulary size but not for overheard speech and later vocabulary size (Shneidman & Goldin-Meadow, 2012; Shneidman et al., 2013; Weisleder & Fernald, 2013), Brown and Gaskins added an important caveat:

However, these studies treat all speech not directly addressed to the child as “overheard,” ignoring the fact that much of the speech (e.g., of adults on the phone, or adult–adult conversations) is irrelevant to the child who may well not be actually “overhearing” it. Such studies need to have more sensitive assessments of what the child is potentially attending to (actually over-hearing) and more subtle analysis of the target vocabulary set in the different settings, before this issue will be clarified. (p. 201)

Brown and Gaskins’ point that speech not addressed to the child should not be equated with

overheard speech is well taken and applies to our current investigation as well.

To address this lacuna, both ethnographic and experimental studies are needed. Ethnographic, observational studies of children in the contexts of everyday life are needed to reveal the kinds of “on-the-ground” attentional stances that become habitual for children when multiparty and bystander configurations are the norm. Such studies will have to grapple with the methodological challenge of determining when children are listening (see Miller et al., 2012). For example, it will be necessary to analyze both participant structures within family conversations as well as topics of talk. In a study of oral narratives in South Baltimore, when young children (2,6) were bystanders to stories told by family members, they were much more likely to contribute relevant verbal responses to stories that focused on their *own* experience, compared with stories about other people’s experiences, implying selective attention to stories about themselves (Miller, Potts, Fung, Hoogstra, & Mintz, 1990). To determine what children actually learn from such experiences, experimental studies, along the lines of those conducted by Rogoff and her team, are also needed. They found that Guatemalan and Mexican-heritage U.S. children, who were present as bystanders and overhearers while another child engaged in an unfamiliar task, attended to what the other child was doing, learned to do the task, and retained what they had learned (Correa-Chávez & Rogoff, 2009; Silva, Correa-Chávez, & Rogoff, 2010). Their attentiveness to the task occurred without any invitation from adults as they waited their turn, in contrast to the European American children, who were less likely to orient to the other child’s activity and to learn how to do the task.

The third important task awaiting future research is to demonstrate how different aspects of communicative competence emerge in relative importance across the preschool years and how the ambient environment supports each in turn. Rowe (2012) demonstrated how even the single language achievement of vocabulary growth is predicted by different aspects of the ambient verbal environment across the preschool years. Yet children also come to be sophisticated storytellers as their language abilities grow. Discourse forms such as narrative flourish in lower income households and are subject to different forms of socialization across groups defined by both culture and income (Miller & Sperry, 2012; Miller et al., 2012; Sperry & Sperry, 1996). Finally, it is commonsense that the relative importance of overheard speech must change as

children grow past the acquisition of early words and come to glean information from teachers and peers in complex multiparty contexts such as the classroom or the playground. Thus, future research must confront the degree to which aspects of appropriate language-learning contexts wax and wane over the course of the early childhood years, including but not limited to the importance of child-directed speech versus overheard speech; the quantity or quality of words in the ambient environment versus the emergence of sophisticated discourse practices; and the relative importance of different interlocutors such as parents, siblings, and peers in the determination of what gets learned by the child.

In addition to these next steps pertaining to variation in young children’s verbal environments and vocabulary learning at home, the last critically important piece of the bigger picture has to do with how this variation relates to what happens when children enter school. Scholars have noted that the practices of nonmajority families are often viewed through a “one-size-fits-all” lens that takes the majority group as the model (Callanan & Waxman, 2013; Cole, 2013; Genishi & Dyson, 2009; Miller & Sperry, 2012; Miller et al., 2005). This paper broadens the lens beyond child-directed speech to include multiparty and bystander speech. It raises the question, “What could be learned about classroom talk in kindergarten and the early grades if the lens were similarly broadened?” Future work will need to address the confluence of vocabulary and discourse in the myriad contexts of children’s everyday lives, the analogous confluence in classrooms, and how the two connect (or not). This approach will offer educators a more nuanced understanding of the full repertoire of verbal means that poor and working-class children bring to the classroom.

References

- Akhtar, N. (2005). The robustness of learning through overhearing. *Developmental Science*, *8*, 199–209. <https://doi.org/10.1111/j.1467-7687.2005.00406.x>
- Akhtar, N., & Gernsbacher, M. A. (2007). Joint attention and vocabulary development: A critical look. *Language and Linguistics Compass*, *1*, 195–207. <https://doi.org/10.1111/j.1749-818X.2007.00014.x>
- Arunachalam, S., Escovar, E., Hansen, M. A., & Waxman, S. R. (2013). Out of sight, but not out of mind: 21-month-olds use syntactic information to learn verbs even in the absence of a corresponding event. *Language and Cognitive Processes*, *28*, 417–425. <https://doi.org/10.1080/01690965.2011.641744>

- Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*, 442–446. <https://doi.org/10.1016/j.cognition.2009.10.015>
- Avineri, N., & Johnson, E. (2015). Introduction. Invited forum: Bridging the “Language Gap.” *Journal of Linguistic Anthropology*, *25*, 67–68. <https://doi.org/10.1111/jola.12071>
- Blum, S. D. (2015). “Wordism”: Is there a teacher in the house? Invited forum: Bridging the “Language Gap.” *Journal of Linguistic Anthropology*, *25*, 74–75. <https://doi.org/10.1111/jola.12071>
- Brown, P., & Gaskins, S. (2014). Language acquisition and language socialization. In N. J. Enfield, P. Kockelman, & J. Sidnell (Eds.), *Cambridge handbook of linguistic anthropology* (pp. 187–226). Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/CBO9781139342872>
- Callanan, M., & Waxman, S. (2013). Commentary on special section: Deficit or difference? Interpreting diverse developmental paths. *Developmental Psychology*, *49*, 80–83. <https://doi.org/10.1037/a0029741>
- Cole, M. (2013). Differences and deficits in psychological research in historical perspective: A commentary on the special section. *Developmental Psychology*, *49*, 84–91. <https://doi.org/10.1037/a0029623>
- Correa-Chávez, M., & Rogoff, B. (2009). Children’s attention to interactions directed to others: Guatemalan Mayan and European American patterns. *Developmental Psychology*, *45*, 630–641. <https://doi.org/10.1037/a0014144>
- Dudley-Marling, C., & Lucas, K. (2009). Pathologizing the language and culture of poor children. *Language Arts*, *86*, 362–370.
- Duncan, G. J., Huston, A. C., & Weisner, T. S. (2008). *Higher ground: New hope for the working poor and their children*. New York, NY: Russell Sage Foundation.
- Duranti, A., Ochs, E., & Schieffelin, B. B. (2012). *The handbook of language socialization*. Malden, MA: Blackwell.
- Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*, 234–248. <https://doi.org/10.1111/desc.12019>
- Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*, 98–116. <https://doi.org/10.1037/0012-1649.42.1.98>
- García Coll, C., & Marks, A. K. (2009). *Immigrant stories: Ethnicity and academics in middle childhood*. Oxford, UK: Oxford University Press.
- Gaskins, S., & Paradise, R. (2010). Learning through observation in daily life. In D. F. Lancy, J. Bock, & S. Gaskins (Eds.), *The anthropology of learning in childhood* (pp. 85–117). Lanham, MD: Alta Mira Press.
- Genishi, C., & Dyson, A. H. (2009). *Children, language, and literacy: Diverse learners in diverse times*. New York, NY: Teachers College Press.
- Goffman, E. (1981). *Forms of talk*. Philadelphia, PA: University of Pennsylvania Press.
- Göncü, A. (Ed.). (1999). *Children’s engagement in the world: Sociocultural perspectives*. Cambridge, UK: Cambridge University Press.
- Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, *35*, 515–531. <https://doi.org/10.1017/S0305000907008641>
- Goodnow, J. J., Miller, P. J., & Kessel, J. (Eds.). (1995). *Cultural practices as contexts for development. New directions for child development*, *67*. San Francisco, CA: Jossey-Bass.
- Hart, B., & Risley, T. R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology*, *28*, 1096–1105. <https://doi.org/10.1037/0012-1649.28.6.1096>
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.
- Hart, B., & Risley, T. R. (1999). *The social world of children learning to talk*. Baltimore, MD: Brookes.
- Hart, B., & Risley, T. R. (2003). The early catastrophe. *Education Review*, *17*, 110–118.
- Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. Cambridge, UK: Cambridge University Press.
- Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–83. <https://doi.org/10.1017/S0140525X0999152X>
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*, 1368–1378. <https://doi.org/10.1111/1467-8624.00612>
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*, 55–88. <https://doi.org/10.1016/j.dr.2005.11.002>
- Hoff, E. (2013). Interpreting the early language trajectories of children from low SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, *49*, 4–14. <https://doi.org/10.1037/a0027238>
- Hudley, E. V. P., Haight, W., & Miller, P. J. (2003). *“Raise up a child”: Human development in an African-American family*. Chicago, IL: Lyceum.
- Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*, F31–F39. <https://doi.org/10.1111/j.1467-7687.2008.00768.x>
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*, 236–248. <https://doi.org/10.1037/0012-1649.27.2.236>
- Huttenlocher, J., Vasilyeva, M., Waterfall, H. R., Vevea, J. L., & Hedges, L. V. (2007). The varieties of speech to

- young children. *Developmental Psychology*, 43, 1062–1083. <https://doi.org/10.1037/0012-1649.43.5.1062>
- Jessor, R., Colby, A., & Shweder, R. (Eds.). (1996). *Ethnography and human development: Context and meaning in social inquiry*. Chicago, IL: University of Chicago Press.
- Johnson, E. J. (2015). Debunking the “language gap.” *Journal for Multicultural Education*, 9, 42–50. <https://doi.org/10.1108/JME-12-2014-0044>
- Kusserow, A. (2004). *American individualisms: Child rearing and social class in three neighborhoods*. New York, NY: Palgrave. <https://doi.org/10.1057/9781403973986>
- Lancy, D. F. (2015). *The anthropology of childhood: Cherubs, chattel, changelings* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Li, J. (2012). *Cultural foundations of learning: East and west*. New York, NY: Cambridge University Press. <https://doi.org/10.1017/CBO9781139028400>
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.
- Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, 11, F9–F16. <https://doi.org/10.1111/j.1467-7687.2008.00671.x>
- Messenger, K., Yuan, S., & Fisher, C. (2015). Learning verb syntax via listening: New evidence from 22-month-olds. *Language Learning and Development*, 11, 356–368. <https://doi.org/10.1080/15475441.2014.978331>
- Miller, P. J. (1982). *Amy, Wendy, and Beth: Learning language in South Baltimore*. Austin, TX: University of Texas Press.
- Miller, P. J., & Cho, G. E. (2018). *Self-esteem in time and place: How American families imagine, enact, and personalize a cultural ideal*. New York, NY: Oxford University Press.
- Miller, P. J., Cho, G. E., & Bracey, J. R. (2005). Working-class children’s experience through the prism of personal storytelling. *Human Development*, 48, 115–135. <https://doi.org/10.1159/000085515>
- Miller, P. J., Fung, H., Lin, S., Chen, E. C.-H., & Boldt, B. R. (2012). How socialization happens on the ground: Narrative practices as alternate socializing pathways in Taiwanese and European-American families. *Monographs of the Society for Research in Child Development*, 77 (1, Serial No. 302), i–140.
- Miller, P. J., Hengst, J. A., & Wang, S.-H. (2003). Ethnographic methods: Applications from developmental cultural psychology. In P. M. Camic, J. E. Rhodes, & L. Yardley (Eds.), *Qualitative research in psychology: Expanding perspectives in methodology and design* (pp. 219–242). Washington, DC: American Psychological Association. <https://doi.org/10.1037/10595-000>
- Miller, P. J., Potts, R., Fung, H., Hoogstra, L., & Mintz, J. (1990). Narrative practices and the social construction of self in childhood. *American Ethnologist*, 17, 292–311. <https://doi.org/10.1525/ae.1990.17.2.02a00060>
- Miller, P. J., & Sperry, D. E. (2012). Déjà vu: The continuing misrecognition of low-income children’s verbal abilities. In S. T. Fiske & H. R. Markus (Eds.), *Facing social class: How societal rank influences interaction* (pp. 109–130). New York, NY: Russell Sage.
- Morgan, K. L. (1980). *Children of strangers: The stories of a black family*. Philadelphia, PA: Temple University Press.
- Neergaard, L. (February 17, 2017). Talk to babies and let them babble back to bridge word gap. *Associated Press: The Big Story*. Retrieved from <http://bigstory.ap.org/article/885621ded7c04a129069cb62eeb2004e/talk-babies-and-let-them-babble-back-bridge-word-gap>
- Ochs, E., & Schieffelin, B. B. (1984). Language acquisition and socialization: Three developmental stories and their implications. In R. A. Shweder & R. A. LeVine (Eds.), *Culture theory: Essays on mind, self, and emotion* (pp. 276–320). Cambridge, UK: Cambridge University Press.
- 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, 763–782.
- Rogoff, B. (2003). *The cultural nature of human development*. New York, NY: Oxford University Press.
- Rogoff, B., Paradise, R., Arauz, R. M., Correa-Chávez, M., & Angelillo, C. (2003). Firsthand learning through intent participation. *Annual Review of Psychology*, 54, 175–203. <https://doi.org/10.1146/annurev.psych.54.101601.145118>
- Rosengren, K. S., Miller, P. J., Gutierrez, I. T., Chow, P. I., Schein, S., & Anderson, K. N. (Eds.). (2014). Children’s understanding of death: Toward a contextualized and integrated account. *Monographs of the Society for Research in Child Development*, 79(1, Serial No. 312), vii–162.
- Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, 35, 185–205. <https://doi.org/10.1017/S0305000907008343>
- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, 83, 1762–1774. <https://doi.org/10.1111/j.1467-8624.2012.01805.x>
- 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: Science and Practice*, 5, 285–310.
- Scott, R. M., & Fisher, C. (2009). 2-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes*, 24, 777–803. <https://doi.org/10.1080/01690960802573236>
- Shneidman, L. A., Arroyo, M. E., Levine, S. C., & Goldin-Meadow, S. (2013). What counts as effective input for word learning? *Journal of Child Language*, 40, 672–686. <https://doi.org/10.1017/S0305000912000141>
- Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, 15, 659–673. <https://doi.org/10.1111/j.1467-7687.2012.01168.x>
- Shneidman, L. A., Sootsman Buresh, J., Shimpi, P. M., Knight-Schwarz, J., & Woodward, A. L. (2009). Social experience, social attention and word learning in an

- overhearing paradigm. *Language Learning and Development*, 5, 266–281. <https://doi.org/10.1080/15475440903001115>
- Shneidman, L. A., & Woodward, A. L. (2016). Are child-directed interactions the cradle of social learning? *Psychological Bulletin*, 142, 1–17. <https://doi.org/10.1037/bul0000023>
- Shweder, R. A., Goodnow, J. J., Hatano, G., LeVine, R. A., Markus, H., & Miller, P. J. (2006). The cultural psychology of development: One mind, many mentalities. In W. Damon (Series Ed.) & R. M. Lerner (Vol. Ed.), *Handbook of child psychology*. Vol. 1: *Theoretical models of human development* (6th ed., pp. 716–792). New York, NY: Wiley.
- Silva, K. G., Correa-Chávez, M., & Rogoff, B. (2010). Mexican-heritage children's attention and learning from interactions directed to others. *Child Development*, 81, 898–912. <https://doi.org/10.1111/j.1467-8624.2010.01441.x>
- Sperry, D. E. (2014). *Listening to all of the words: Reassessing the verbal environments of young working-class and poor children*. Unpublished doctoral dissertation, University of Illinois, Urbana–Champaign, IL.
- Sperry, L. L., & Sperry, D. E. (1996). Early development of narrative skills. *Cognitive Development*, 11, 443–465. [https://doi.org/10.1016/S0885-2014\(96\)90013-1](https://doi.org/10.1016/S0885-2014(96)90013-1)
- Sperry, L. L., & Sperry, D. E. (2000). Verbal and nonverbal contributions to early representation: Evidence from African American toddlers. In N. Budwig, I. Č. Užgiris, & J. V. Wertsch (Eds.), *Communication: An arena of development* (pp. 143–165). Stamford, CT: Ablex.
- Sperry, D. E., & Sperry, L. L. (2016). Counting in context: Studying children's everyday talk by combining numbers and words. In J. Prior & J. Van Herwegen (Eds.), *Practical research with children* (pp. 209–227). Oxford, UK: Routledge.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Erlbaum.
- Ward, M. C. (1971). *Them children: A study in language learning*. New York, NY: Holt, Rinehart and Winston.
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24, 2143–2152. <https://doi.org/10.1177/0956797613488145>
- Weisner, T. S. (Ed.). (2005). *Discovering successful pathways in children's development: Mixed methods in the study of childhood and family life*. Chicago, IL: University of Chicago Press.
- Wiley, A. R., Rose, A. J., Burger, L. K., & Miller, P. J. (1998). Constructing autonomous selves through narrative practices: A comparative study of working-class and middle-class families. *Child Development*, 69, 833–847. <https://doi.org/10.1111/j.1467-8624.1998.00833.x>
- Wolcott, H. F. (1995). *The art of fieldwork*. Walnut Creek, CA: AltaMira Press.
- Yuan, S., & Fisher, C. (2009). “Really? She blicked the baby?” Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, 20, 619–626. <https://doi.org/10.1111/j.1467-9280.2009.02341.x>
- Zentella, A. C. (2015). Books as the magic bullet. Invited Forum: Bridging the “Language Gap.” *Journal of Linguistic Anthropology*, 25, 75–77. <https://doi.org/10.1111/jola.12071>