Received: 30 October 2019 | Accepted: 29 October 2020

DOI: 10.1111/infa.12378

INFANCY
THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES
WILEY

RESEARCH ARTICLE

Beyond talk: Contributions of quantity and quality of communication to language success across socioeconomic strata

Lillian R. Masek[1] | **Sarah J. Paterson**[2] | **Roberta Michnick Golinkoff**[3] | **Roger Bakeman**[4] | **Lauren B. Adamson**[4] | **Margaret Tresch Owen**[5] | **Amy Pace**[6] | **Kathy Hirsh-Pasek**[1]

[1]Department of Psychology, Temple University, Philadelphia, PA, USA

[2]James S. McDonnell Foundation, Saint Louis, MO, USA

[3]School of Education, The University of Delaware, Newark, DE, USA

[4]Department of Psychology, Georgia State University, Atlanta, GA, USA

[5]School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, TX, USA

[6]Department of Speech and Hearing Sciences, The University of Washington, Seattle, WA, USA

Correspondence
Lillian R. Masek, Department of Psychology, Temple University, 1701 N. 13th St., 6th Floor, Weiss Hall Philadelphia, PA 19122, USA.
Email: lillian.masek@temple.edu

Funding information
Bezos Family Foundation, Grant/Award Number: 270642 18110-02; Institute of Education Sciences, Grant/Award Number: R305A110284 and R324A160241; Temple University, Grant/Award Number: University Fellowship; William Penn Foundation, Grant/Award Number: 45-15

Abstract

Infants from low-socioeconomic status (SES) households hear a projected 30 million fewer words than their higher-SES peers. In a recent study, Hirsh-Pasek et al. (*Psychological Science*, 2015; 26: 1071) found that in a low-income sample, *fluency and connectedness* in exchanges between caregivers and toddlers predicted child language a year later over and above quantity of talk (Hirsh-Pasek et al., *Psychological Science*, 2015; 26: 1071). Here, we expand upon this study by examining *fluency and connectedness* in two higher-SES samples. Using data from the NICHD Study of Early Child Care and Youth Development, we sampled 20 toddlers who had low, average, and high language outcomes at 36 months from each of 2 groups based on income-to-needs ratio (INR; middle and high) and applied new coding to the mother–toddler interaction at 24 months. In the high-INR group, the quality of mother–toddler interaction at 24 months accounted for more variability in language outcomes a year later than did quantity of talk, quality of talk, or sensitive parenting. These results could not be accounted for by child language ability at 24 months. These effects were not found in the middle-INR sample. Our findings suggest that when the

quality of interaction, *fluency and connectedness*, predicts language outcomes, it is a robust relation, but it may not be universal.

1 | INTRODUCTION

By school entry, many children from low-socioeconomic status (SES) backgrounds trail behind their higher-SES peers in both language comprehension and production (Ginsborg, 2006; Hart & Risley, 1995; Levine et al., 2020). These early language abilities predict school readiness levels at the start of school and academic skills throughout the elementary school years (Pace et al., 2019). Thus, children from low-SES backgrounds are often at a higher risk for poor academic achievement (Hoff, 2013).

Hart and Risley's (1995) seminal study on SES and language development identified early differences in infant-directed speech and projected that, by age 4, children in poverty heard 30 million fewer words than their more affluent peers. This word gap translated into differences in the children's vocabulary and IQ (Hart & Risley, 1995). Hart and Risley's study has sparked rigorous debate (Golinkoff et al., 2019; Purpura, 2019; Sperry et al., 2019a), and its findings should be interpreted with caution. It has been criticized for its small sample size (Pan et al., 2005), the homogeneity of its low-SES sample (Dudley-Marling & Lucas, 2009; Sperry et al., 2019b), and for how talk was measured (Michaels, 2013; Sperry et al., 2019b). Further, researchers have worried about how SES was defined in the Hart and Risley study and subsequent research. SES is a multifaceted construct and, as such, is operationalized in a variety of ways across different studies, including income, education, and use of government assistance. Despite these criticisms, however, the finding that young children from lower-SES backgrounds hear less talk has been widely replicated across operationalizations of the SES construct (e.g., Cartmill et al., 2013; Gilkerson et al., 2017; Hoff, 2003; Huttenlocher et al., 2010; Levine et al., 2020; Vanormelingen & Gillis, 2016).

While the findings on quantity are debated, few debate the other key point that Hart and Risley made about the importance of the *quality* of the early language environment. They wrote, "Not all talk is equally informative, and not all parent–child interactions are equally encouraging" (Hart & Risley, 1995, p. 95). The quality of the early language environment is measured in a variety of ways, including global factors such as parental sensitivity and cognitive stimulation (Leigh et al., 2011; Mistry et al., 2008; Raviv et al., 2004) and interactional factors such as lexical and syntactic diversity (Hoff, 2003; Huttenlocher et al., 2010; Rowe, 2012). Like quantity of talk, studies have found SES-related differences across many facets of quality (Hoff, 2003; Huttenlocher et al., 2010; Mistry et al., 2008; Raviv et al., 2004). Studies that examine SES differences in quantity and quality of language input frequently conclude that variation in these features of the early language environment account for the observed SES-related differences in children's own language use (see Luo et al., 2016 for a review).

Despite these findings, SES differences are not found in all aspects of language input. Cartmill and colleagues found that although low-SES families talked less to their children overall, there were no SES differences in the referential transparency of the speech, or how well the meaning of a word could be inferred from its context, which was a robust predictor of children's later language skills (Cartmill et al., 2013). Furthermore, several studies that find SES differences also show large variability *within* SES strata (Cartmill et al., 2013; Gilkerson et al., 2017; Huttenlocher et al., 2010; Pan et al., 2005). Gilkerson et al. (2017) found that the average amount of verbal engagement, as measured by conversational turns, by the most talkative half of the lowest-SES sample was higher than that of the less talkative half of the highest-SES sample, suggesting that many lower-SES parents are talking and engaging in more conversation than their higher-SES counterparts.

In all these studies on SES differences, SES and language outcomes are closely related, making it difficult to tease apart SES differences in language environments that have implications for later outcomes and those that do not. Furthermore, it is difficult to determine whether the mechanisms for language development are the same across SES groups, despite differences in the language environment. Here, using income-to-needs ratio (INR, defined subsequently) as an approximate proxy for SES, we attempt to shed light on these observed differences. Drawing upon a large, archived project, we examined predictors of language outcomes in children from low-, middle-, and high-INR groups. In each of these three groups, we selected equal numbers of boys and girls and equal numbers of children with low, middle, and high expressive language outcome. Thus, for each INR group, we could examine the full range of possible outcomes one year after we observed the quality of the early language environment and early caregiver–child interactions.

1.1 | A foundation for communication

High-quality interactions between a supportive conversational partner, the child, and a shared object or event are optimal for children to learn new linguistic items (Adamson et al., 2004; Tamis-LeMonda et al., 2014; Tomasello & Farrar, 1986). Neuroscience research suggests that engaging children in back-and-forth conversations (though not the number of words heard) relates to development in regions of the brain related to language (Romeo, Leonard, et al., 2018; Romeo, Segaran, et al., 2018). In a study of a low-INR sample by Hirsh-Pasek et al. (2015), caregiver–toddler interaction defined by dyadic conversations infused with symbols, scripted patterns of behavior, and fluid conversation, the so-called *communication foundation*, was a stronger predictor of toddler language outcomes than was the number of words the caregiver said.

The current study extends the findings from Hirsh-Pasek et al. (2015)—which used the systematic observational rating methods detailed in Adamson et al. (2016) to describe the *communication foundation* that facilitates language development—to two new groups. Hirsh-Pasek et al. examined this construct in a low-INR sample (income-to-needs ratio less than 1.8, as explained subsequently) and found that how well mother–toddler dyads worked together to build this foundation at 24 months predicted expressive language at 36 months, above and beyond the effects of the quantity of language toddlers heard. These findings held even when sensitive parenting was controlled for. The rating item that accounted for the largest amount of variance in their study was called *fluency and connectedness*, the smooth, back-and-forth communication, both verbal and nonverbal, between caregiver and toddler. Here, we examine *fluency and connectedness* across and within two new groups with higher income-to-needs ratios (middle and high INR) that, like Hirsh-Pasek et al., contain equal numbers of children with low, middle, and high 36-month expressive language scores.

Although *fluency and connectedness* has been investigated in diverse samples, there have been mixed results as to its relation to language outcomes. Some studies find a relation between *fluency and connectedness* and later language (Conway et al., 2018; Smith et al., 2018), and others find a relation in some groups but not all (Adamson et al., 2019). Given the diversity of findings, it is essential to examine this construct across varying levels of household income among children who are diverse in terms of language outcomes. If *fluency and connectedness* predicts language outcomes in our middle-INR and high-INR samples, it would suggest that it is a central context for early language learning such that regardless of income level, a toddler who is engaged in more, higher-quality, fluid-and-connected interactions early in development will have better language skills later in development. If *fluency and connectedness* does not predict language outcomes in our middle-INR and high-INR samples, it would suggest that it is more limited in its relation to later language outcomes such that *fluency and connectedness* is more important for low-INR toddlers (Hirsh-Pasek et al., 2015) and less important for middle- and high-INR toddlers. If this is found, it could be because children from families with more income have more access to learning materials (e.g., books) and experiences (e.g., museums) that promote language skills outside of the context of the caregiver–child interaction (Pace et al., 2017).

It is also likely that *fluency and connectedness* relates to other measures of quality language input such as the diversity of words toddlers hear, though this has yet to be determined. Given that both sensitive parenting and quality of language input vary as a function of income level and relate to later language, it is important to determine whether *fluency and connectedness* relates to language outcomes above and beyond these other measures. If *fluency and connectedness* predicts language outcomes across income levels, and these effects cannot be accounted for solely by either sensitive parenting or quality of language input, then we have isolated a component of communication that is likely critical in language development during toddlerhood.

1.2 | Current Study

This study examines parent–toddler interaction at 24 months of age and its relation to child expressive language at 36 months of age. Extending Hirsh-Pasek et al. (2015), who focused on a low-INR sample, we likewise sampled from the NICHD Study of Early Child Care and Youth Development (NICHD SECCYD) archive, selecting toddlers from families with middle and high levels of INR. Like Hirsh-Pasek et al., within each INR group we selected equal numbers of boys and girls and equal numbers of children who scored low, average, and high on 36-month expressive language. The present study compares data from the low-INR group from Hirsh-Pasek et al. with our middle- and high-INR groups. This allowed us to compare groups with different income levels, each of which contained equivalent ranges of 36-month expressive language variability, as we examined differences in early language environments and probed whether the *fluency and connectedness* rating captured a general feature of language development within each INR group, at least in this sample from the United States. To provide comparison, we included both 36-month expressive and receptive language in our analyses (Hirsh-Pasek et al. only included expressive language). Expressive and receptive language are considered by some to represent separable skills that, although related, are not equivalent, especially in toddlers. For example, among children with specific language impairment, most have impairment in either expressive or receptive language (e.g., Law et al., 2000).

By comparing language-equivalent samples of toddlers from low-, middle-, and high-INR groups, we sought to answer the following questions:

1. Is the *fluency and connectedness* rating an equally potent predictor of child language outcomes across language-equated INR groups? We hypothesize that the mechanisms that spark language learning are grounded in *fluency and connectedness* and as such, for all three INR groups, *fluency and connectedness* at 24 months will relate to child language skills at 36 months.
2. Is the *fluency and connectedness* rating a unique predictor of child language outcomes across language-equated INR groups when accounting for quantity of language input, quality of language input in terms of lexical diversity, and sensitive parenting? We hypothesize that *fluency and connectedness* will relate to child language skills at 36 months even when these other variables are controlled.

3. Across INR groups, is the *fluency and connectedness* rating a unique predictor of child language outcomes when accounting for earlier language skills? We hypothesize that not only does *fluency and connectedness* relate to child language skill, but that it will continue to predict language skills at 36 months when controlling for language at 24 months.

2 | METHOD

This study is an extension of Hirsh-Pasek et al. (2015) and, like it, draws data and uses video records from the archived NICHD SECCYD, a large-scale, longitudinal study on the effects of childcare on cognitive, language, and socio-emotional development as well as health (see https://www.nichd.nih.gov/research/supported/seccyd). Data for the NICHD SECCYD were collected from 10 sites across the United States and represented the demographics of the United States at the time. Participants were recruited in 1991 and followed from birth to age 15 years. Several publications from this study document the recruitment and demographics as well as details of the study's methodology (e.g., NICHD Early Child Care Research Network, 2002, 2006).

The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. All procedures involving human subjects in this study were approved by the Temple University Internal Review Board (protocol number 21978).

2.1 | Sample selection

Hirsh-Pasek et al. (2015) reported on a low-INR sample (income-to-needs ratio less than 1.8). To assure equal representation of children with low, medium, and high expressive language ability, they selected 20 children each (10 boys and 10 girls) from the lower third, middle third, and highest third of eligible children based on their Reynell 36-month expressive language scores (Reynell, 1991; a standardized language outcome measure). We used the same strategy to select a middle-INR and high-INR sample (income-to-needs ratio 1.8–4.0 and greater than 4.0, respectively). Thus, our analyses could compare three samples—one from Hirsh-Pasek et al.—each with an equivalent distribution of expressive language ability.

Like Hirsh-Pasek et al. (2015), when selecting the middle- and high-INR samples, we considered cases with a valid income-to-needs ratio score when children were 24 months old, a valid expressive language score on the Reynell at 36 months, and a usable video recording of interaction at 24 months and selected cases from 5 of the study's 10 sites—Pittsburgh, Pennsylvania; Morganton, North Carolina; Charlottesville, Virginia; Philadelphia, Pennsylvania; and Madison, Wisconsin—because these sites generally had better video recordings than some of the other sites. When selecting the low-INR sample, Hirsh-Pasek et al. oversampled children with higher expressive language scores. Similarly, when selecting the high-INR sample, we oversampled children with lower expressive language scores. Oversampling was not required for the middle-INR group (see Table 1 for details on sample selection). The mean expressive language scores were 100 [96, 104], 99 [95,103], and 98 [95,102] for the low-, middle-, and high-INR groups (95% confidence intervals in brackets), which confirms that our selection strategy resulted in equivalent language score distributions.

TABLE 1 Number of valid cases in the full, selection, and study samples

INR group	Low tertile	Mid tertile	High tertile
Sample	***N* (boys, girls)**	***N* (boys, girls)**	***N* (boys, girls)**
Middle			
Full sample	128 (80, 48)	174 (83, 91)	113 (49, 64)
Selection sample	52 (32, 20)	89 (37, 52)	64 (28, 36)
Study sample	20 (10, 10)	20 (10, 10)	20 (10, 10)
High			
Full sample	86 (56, 30)	168 (77, 91)	109 (44, 65)
Selection sample	34 (21, 13)	74 (35, 39)	57 (24, 33)
Study sample	20 (10, 10)	20 (10, 10)	20 (10, 10)

Note: The full sample included cases with valid expressive language and income-to-needs ratio scores with usable video. The selection sample is limited to the five sites from which study sample cases were selected. The study sample consists of the participants used in this study. For sample selection of the low-INR sample, see Hirsh-Pasek et al., 2015.

2.2 | Sample characteristics

Children's mean ages at the 24-month visit for the Hirsh-Pasek et al. (2015) low-INR sample and for our two new samples were 25.2, 25.3, and 25.2 months (*SD*s =1.02, 0.88, and 0.87), respectively. By design, age did not differ by INR group; one-way analysis of variance $\eta^2 = .004$, $p = 72$. Household income, mother's education, and proportion White, however, did differ. As would be expected, selecting groups by INR, the distribution of reported household incomes hardly overlapped (see Figure 1, top): 92% of the household incomes for the low-INR group were less than $30,000, 87% for the middle-INR group were between $30,00 and $59,999, and 83% in the high-INR group were higher than $59,999. These are 1993 dollars. In terms of 2020 dollars, these amounts would be almost doubled (to $57,500 and $115,000, respectively; bureau of labor statistics, https://data.bls.gov for comparison, corresponding 2018 percentiles were 46% and 75%; https://www.census.gov).

The distribution of mothers' years of education overlapped a bit more (see Figure 1, bottom), but still differed by INR group: 92% of mothers in the low-INR group reported no education past 14 years; 90% in the middle-INR group reported no education past 16 years, whereas 65% in the high-INR group reported 16 or more years of education. Finally, percentages White were 63%, 80%, and 93% for the low-, middle, and high-INR groups, respectively (all chi-squared were significant, $p < .001$; p values for odds ratios comparing low- to middle-INR groups and middle- to high-INR groups were all less than .05).

In Hirsh-Pasek et al. (2015), the low-INR group was selected to minimize differences in maternal education and child race across child language ability. Similarly, for the middle- and high-INR samples, there were no significant differences in either maternal education or child race across child language ability tertiles (all chi-squared were nonsignificant, $p > .05$).

2.3 | The three boxes mother–child interaction task

As described in Hirsh-Pasek et al. (2015), part of the SECCYD 24-month laboratory visit included a 15-minute semi-structured play session with mothers and their toddlers. The interaction was video-recorded through a one-way mirror, and sound was recorded using a microphone in the room. Mothers

Distributions of household income and years of maternal education

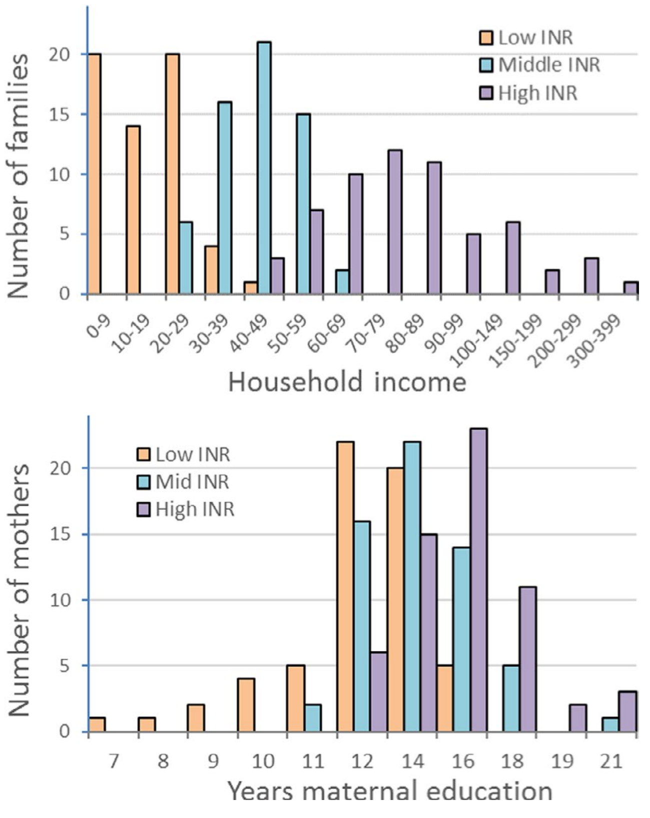

FIGURE 1 Distributions of household income and years of maternal education

were instructed to play with their toddler using the contents of three boxes (labeled 1, 2, and 3), beginning with the first box and ending with the third box. The first box contained a picture book, *Barnyard Tracks*. The second box contained a toy kitchen with a shape-sorter, and the third box included an activity house with a puppy, a baby, and a car. Mothers were given no instruction on how much of the 15 minutes to spend with each box and, if they asked, were told to handle the time with their toddler and the boxes however they liked. All mothers spoke English and used English during the interaction.

2.4 | Measures of the interaction between parent and child

Hirsh-Pasek et al. (2015) derived variables from observations of these video records. We did the same for the middle- and high-INR samples. As detailed in Hirsh-Pasek et al., the measure of *fluency and connectedness* was rated from the 24-month video records.

2.4.1 | Fluency and connectedness

The *fluency and connectedness* rating measures the balance of the communication between mother and toddler as well as the ease of that communication. Dyads were considered to display high *fluency* *and connectedness* if both mother and toddler took turns initiating interaction and if mother and toddler took conversational turns without overlap or long gaps in communication. Conversely, the *fluency and connectedness* score was considered poor if no conversation was established (i.e., bids for conversation went unanswered), or if one partner (either the mother or the child) dominated the conversation. It is important to note that conversation in this context was not necessarily verbal. If the mother asked, "where's the doggy?" and the toddler held up the doggy, that would be considered a conversational exchange.

The *fluency and connectedness* item was rated on a scale from 1 to 7 considering both the number of episodes and the quality of the conversation (Adamson et al., 2016). An example of a score of 1 would be a dyad who engaged in no conversation at all during the observation. An example of a score of 4 would be either a dyad who engaged in a few, balanced, and smooth conversational bouts or in many bouts dominated by the mother. An example of a score of 7 would be a dyad who engaged in smooth, balanced conversation throughout the interaction.

Training for the low-INR group raters was described in Hirsh-Pasek et al. (2015). For the middle- and high-INR group, raters received a month of training on six items of the Joint Engagement Rating Inventory (Adamson et al., 2016) before they rated the video records. This training included a day-long workshop led by researchers who had extensive experience with the rating items used to characterize parent–toddler interactions and had coded the low-INR videos from Hirsh-Pasek et al. (2015). Raters used a rating manual (Adamson et al., 2016) that detailed each rating item. Rating typically entailed viewing the video record of a session three times, during which the rater took notes related to interactional behaviors including *fluency and connectedness*, the item analyzed in this study. Raters were blind to other information about the toddlers, including their Reynell score, except for observed and inferred racial differences.

To check agreement for the middle- and high-INR ratings, 12 of the 60 sessions (20%) in each INR group were coded by one of two reliability raters. Both of the reliability raters had served as raters for the low-INR sample from Hirsh-Pasek et al. (2015). The primary rater for the middle- and high-INR groups did not know which sessions were to be double-rated. Weighted kappas (Bakeman & Quera, 2011)—weighting 1-point disagreements 0 (i.e., effectively considering them agreements), weighting 2-point disagreements 1, 3-point disagreements 2, etc.—were .89 and .68 for the middle-, and high-INR groups, respectively. To produce kappas of this magnitude for these 1–7 weighted ratings with these marginals, simulated observers would need to be at least 95% accurate on average (computed by the KappaAcc program, Bakeman, 2018).

2.4.2 | Talk variables

Additionally, we derived variables from transcripts we made of these video records. We transcribed the three boxes sessions for the middle- and high-INR groups using the Codes for Human Analysis of Transcripts (CHAT) conventions and then analyzed them using Computerized Language Analysis (CLAN; MacWhinney, 2000). All transcripts were reviewed for accuracy, and any discrepancies were settled by a third party. CLAN was used because it allows for the computation of VocD, a measure of lexical diversity described subsequently. Hirsh-Pasek et al. (2015) used the Systematic Analysis of Language Transcripts program (SALT; Miller & Iglesias, 2010). To derive VocD, and stay consistent with the middle- and high-INR samples, the transcripts from Hirsh-Pasek et al. (2015) were converted to the CLAN format for analysis in this study. Consequently, values we report for the quantity of talk in the low-INR group vary slightly from those given in Hirsh-Pasek et al. The three language variables we use for this report are:

1. *Quantity of maternal talk*. Quantity of the mother's talk during the interaction was assessed by maternal words per minute (*WPM*). Words per minute measures the number of words the mother uses divided by the length of the interaction in minutes. This measure is used as a measure of the total amount of talk the child was exposed to that controls for the length of the interaction.
2. *Lexical diversity of maternal talk*. Lexical diversity of the mother's talk during the interaction was measured using the *VocD* algorithm (Malvern et al., 2004). VocD scores are computed by randomly selecting portions of a person's speech and calculating the type-to-token ratio for each section, then averaging them together. The average number of utterances for each group were 250 for the low-INR group (range =71–418), 288 for the middle-INR group (range =113–528), and 298 for the high-INR group (range =127–500). This method of calculating lexical diversity was used because it does a better job of controlling for the overall amount of talk than the number of types per minute and a better job of controlling for length of the interaction than the standard type-to-token ratio.
3. *Quantity of child talk*. Child talk during the interaction, like maternal talk, was assessed by words per minute (*WPM*). Words per minute measures the number of words a toddler uses divided by the length of the interaction in minutes. This measure is used as a proxy for the toddler's vocabulary size and overall talkativeness that controls for the length of the interaction.

2.5 | Additional variables

Additionally, we used two variables from the SECCYD archive, sensitive parenting and child's 24-month expressive language. Sensitive parenting had been used previously in the Early Child Care Research Network's report addressing childcare effect sizes for different developmental domains (NICHD Early Child Care Research Network, 2006) and was also used by Hirsh-Pasek et al. (2015) for the low-INR sample.

2.5.1 | Sensitive parenting

As described in Hirsh-Pasek et al. (2015), sensitive parenting is a composite, created by standardizing and averaging SECCYD scores for maternal sensitivity and stimulation and responsiveness of the family environment. Maternal sensitivity was the sum of three rating items—sensitivity and responsiveness to the toddler, positive regard, and intrusiveness (reverse scored)—made from video records of the mother–child interaction and averaged over three ages (6, 15, and 24 months). Stimulation and responsiveness of the family environment was assessed with the total score of the Home Observation for Measurement of the Environment (HOME; Caldwell & Bradley, 1984)—a measure that taps into high-quality aspects of the environment such as variety in daily activities and includes parenting actions such as acceptance of child's behavior—averaged over 15 and 36 months because the HOME was not administered at 24 months. Details of these procedures and the validity and reliability of the interaction procedures and ratings of maternal sensitivity and of the HOME are well documented (NICHD Early Child Care Research Network, 1999, 2002).

2.5.2 | Child's 24-month expressive language

To provide a second measure of the child's 24-month expressive language ability, independent of the WPM derived from the transcripts, we took the number of words the child was reported to produce at

TABLE 2 Descriptive and analysis of variance statistics for study variables

	Language-equated INR group						ANOVA statistics		Cohen's *d*		
	Low		Middle		High						
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	η^2	*p*	Low–middle	Low–high	Middle–high
Maternal WPM	51_a	18	57_{ab}	18	62_b	16	.069	.002	0.67	0.37	0.29
Maternal lexical diversity	58_a	14	63_b	13	69_c	11	.134	<.001	0.97	0.41	0.55
Sensitive parenting	-0.51_a	0.87	0.13_b	0.62	0.40_b	0.57	.232	<.001	1.23	0.85	0.44
Child WPM	7.8_a	5.0	11.9_b	6.1	12.5_b	6.4	.114	<.001	0.82	0.73	0.11
CDI word count	379	169	354	173	354	155	.005	.661	–0.15	–0.15	0.002
Fluency and connectedness	3.85	1.13	3.97	1.16	4.20	1.20	.016	.250	0.30	0.10	0.20
Child expressive language[a]	100	15	99	16	98	14	.002	.874	–0.10	–0.06	–0.04
Child receptive language[a]	95	15	99	12	106	14	.102	<.001	0.77	0.31	0.53

Note: $n = 60$ for each INR group except for the CDI score, for which $n = 53$, 55, and 57 for the low-, middle-, and high-INR group, respectively. All variables were assessed at 24 months except for the child language scores at 36 months, as noted. η^2 and *p* values are from a one-way analysis of variance. *Fluency and connectedness* was rated on a scale of 1–7. Means that do not differ $p < .05$ per a Tukey HSD post hoc test share a common subscript. Except for maternal lexical diversity and child receptive language, other means for the low-INR group were reported in Hirsh-Pasek et al. (2015).

[a]Assessed at 36 months.

24 months from the SECCYD archive. This number, which approximates the toddler's productive vocabulary size, had been assessed with the MacArthur Communicative Development Inventory (CDI): Words and Sentences (now called the MacArthur-Bates Communicative Development Inventory or MB-CDI; Fenson et al., 1991). Some CDI scores were missing: 7 from the low-, 5 from the middle-, and 3 from the high-INR groups.

2.6 | Measures of language outcome

We used 36-month expressive and receptive language scores from the SECCYD archive as assessed with the Reynell Developmental Language Scales (Reynell, 1991). The Reynell consists of two 67-item subscales: expressive language and verbal comprehension. The expressive language subscale assesses how well children can define words as well as label and describe objects and activities from picture. The verbal comprehension subscale assesses how well children can identify objects and follow verbal instruction. These subscales are norm-referenced for children age 1–7 years and were chosen for their high validity (Reynell, 1991). Hirsh-Pasek et al. (2015) used only the expressive language score, but for this study we use both to see whether the relations are equivalent for expressive and receptive language.

3 | RESULTS

Before addressing our research questions, we ran analyses for descriptive purposes. First, we examined whether the means for each of our study variables differed across INR groups. Results are given in Table 2. Low-INR group means were lower than middle-INR group means, which in turn were lower than high-INR group means, with just two exceptions: As a result of our sampling strategy, 36-month expressive language scores were essentially the same in all three groups, and CDI counts were greater, but only weakly and not significantly so, in the low-INR group (Cohen's *d*s comparing low- to middle- and high-INR groups were both –.15; thresholds for small or weak, medium or moderate, and large or strong *d*s are 0.20, 0.50, and 0.80; Cohen, 1988), Otherwise—in order of magnitude for statistically significant differences—low-INR means were strongly lower than high-INR means for sensitive parenting, maternal lexical diversity, and child WPM; and moderately lower for child's 36-month receptive language and maternal WPM. Low-INR means were strongly lower than middle-INR means for sensitive parenting, moderately lower for child WPM, and weakly lower for maternal lexical diversity. And middle-INR means were moderately lower than high-INR means for maternal lexical diversity and child's 36-month receptive language. For exact *d* values, see Table 2.

Next, we examined correlations between our study variables, first for the whole sample Table 3) and then separately for the three INR groups (Table 4). For the whole group, of 28 correlations, 32% were small, 61% medium, and 7% large (thresholds for small, medium, and large are .10, .30, and .50 absolute; Cohen, 1988); $p < .05$ for all. Both the magnitude of the correlations and their statistical significance varied by INR group. In the low-INR group, 11% were small, 64% were medium, 25% were large, and none failed to meet the threshold for small; $p < .05$ for all. In contrast, in the middle-INR group 36% were small, 25% were medium, 14% were large, and 25% failed to meet the threshold for small; $p < .05$ for 54%. In the high-INR group 43% were small, 39% were medium, 14% were large, and 4% failed to meet the threshold for small; $p < .05$ for 75%. All correlations were positive (except for one that failed to meet the threshold for small). Of the 28 correlations, six differed significantly between low- and middle-INR groups, one between low- and high-INR groups, and two between

TABLE 3 Correlations for study variables: whole sample

Variable	mWPM	VocD	SensP.	cWPM	CDI	F&C	ExpLang
Maternal WPM	—						
Maternal lexical diversity	.45	—					
Sensitive parenting	.47	.46	—				
Child WPM	.21	.25	.33	—			
Child CDI (n = 165)	.18	.18	.22	.46	—		
Fluency and connectedness	.36	.33	.36	.65	.47	—	
Child expressive language[a]	.17	.25	.17	.28	.40	.33	—
Child receptive language[a]	.31	.45	.50	.43	.37	.39	.46

Note: $N = 180$ unless otherwise indicated. All variables were assessed at 24 months except for the child language scores at 36 months, as noted. All correlations were significant $p < .05$.

[a]Assessed at 36 months.

middle- and high-INR groups (see Table 4). Partly these differences reflect one anomalous case, a girl in the middle-INR group who had the highest 36-month expressive language score but whose WPM score was among the five lowest and whose CDI word count was among the six lowest in her group, and who was rated two on *fluency and connectedness* (no one scored lower in her group). Without her data, five of the 28 correlations would differ significantly between low- and middle-INR groups, still one between low- and high-INR groups, but none between middle- and high-INR groups.

It is noteworthy that all associations between our predictor variables and children's 36-month expressive and receptive language scores were positive (albeit one near zero) and were for each INR group:

1. moderate or strong for the four mother and child language variables, sensitive parenting, and *fluency and connectedness* in the low-INR group;
2. moderate only for children's CDI count, but otherwise weak or near zero for the other variables in the middle-INR group;
3. likewise moderate or strong in the high-INR group for the same variables listed for the low-INR group, except weak for maternal WPM and hardly at all (near zero) for sensitive parenting with expressive language.

3.1 | Effect of fluency and connectedness on language outcomes across INR groups

To address our first research question—whether *fluency and connectedness* was an equally potent predictor of 36-month child language outcomes in all three language-equated INR groups—we considered the correlations given in Table 4. The answer is no. *Fluency and connectedness* correlated strongly or near strongly with both the child's 36-month expressive and receptive language in the low- and high-INR groups but not in the middle-INR group. For expressive language, the just strong

TABLE 4 Correlations for study variables separately by INR group

Variable	mWPM	VocD	SensP.	cWPM	CDI	F&C	ExpLang
Low-INR group							
Maternal WPM	—						
Maternal lexical diversity	**.52**	—					
Sensitive parenting	**.42**	**.29**	—				
Child WPM	**.34**$_a$	**.37**$_a$	**.37**	—			
Child CDI ($n = 53$)	**.30**	**.29**	**.31**	**.47**	—		
Fluency and connectedness	**.48**	**.51**$_a$	**.37**	**.75**	**.48**	—	
Child expressive language[a]	**.34**	**.40**	**.35**$_a$	**.48**$_a$	**.43**	**.52**$_a$	
Child receptive language[a]	**.34**	**.52**	**.59**$_a$	**.45**	**.35**	**.41**$_{ab}$	**.52**
Middle-INR group							
Maternal WPM	—						
Maternal VocD	**.35**	—					
Sensitive parenting	**.47**	**.52**	—				
Child WPM	−.09$_b$	−.06$_b$	.05	—			
Child CDI ($n = 55$)	.17	.09	**.35**	**.50**	—		
Fluency and connectedness	**.28**	.13	**.41**	**.56**	**.49**	—	
Child expressive language[a]	.05	.16	**.26**$_{ab}$	.13$_b$	**.34**	.06$_b$	
Child receptive language[a]	.03	**.26**	**.29**$_b$	.22	**.35**	.17$_a$	**.57**
High-INR group							
Maternal WPM	—						
Maternal VocD	**.30**	—					
Sensitive parenting	**.33**	**.27**	—				
Child WPM	.21$_{ab}$	.22$_{ab}$	.19	—			
Child CDI ($n = 57$)	.16	**.33**	.20	**.58**	—		
Fluency and connectedness	**.26**	**.29**$_{ab}$	**.30**	**.69**	**.48**	—	
Child expressive language[a]	.18	**.30**	−.04$_b$	**.35**$_{ab}$	**.43**	**.45**$_a$	
Child receptive language[a]	**.37**	**.30**	**.33**$_{ab}$	**.46**	**.58**	**.51**$_b$	**.40**

Note: $N = 60$ for each INR group unless otherwise indicated. All variables were assessed at 24 months except for the child language scores at 36 months, as noted. Correlations significant $p < .05$ are bolded. Correlations that differ between groups $p < .05$ per Fisher r to z test have different subscripts. Except for maternal lexical diversity and child receptive language, other correlations for the low-INR group were reported in Hirsh-Pasek et al. (2015).

[a]Assessed at 36 months.

correlation in the low-INR group (.52) and almost strong correlation in the high-INR group (.45) differed significantly from the less-than-small correlation in the middle-INR group (.06). For receptive language, the small correlation in the middle-INR group (.17) differed significantly from the just strong correlation in the high-INR group (.51), but not from the moderate correlation in the low-INR group (.41).

3.2 | Unique effects of fluency and connectedness on language outcomes

To address our second and third research questions—whether *fluency and connectedness* was a unique predictor of 36-month child language outcomes across all three language-equated INR groups, above and beyond the quantity of language input (maternal WPM), quality of language input (maternal lexical diversity), and sensitive parenting (first question) and above and beyond children's WPM and CDI counts (second question)—we employed regression models. We regressed each 36-month language score first on one of the five predictor variables (step 1) and then added *fluency and connectedness* (step 2). This gives the total variance accounted for by each two-variable model (total R^2) and the unique variance accounted for by *fluency and connectedness* additionally (ΔR^2). Reversing the order—entering *fluency and connectedness* first—gives the unique variance accounted for by each of the five predictor variables. Finally, subtracting the two unique portions of variance from the total R^2 gives the overlap—the proportion of total variance accounted for by the two-predictor variables jointly, as in a Venn diagram (see Cohen & Cohen, 1983).

The results for children's 36-month expressive language are shown graphically in Figure 2 and for children's 36-month receptive language in Figure 3. The figures are stacked bar graphs, with each bar divided into three parts. The height indicates the total R^2 for the particular two-variable model. The top portion represents the ΔR^2 for the indicated variable when added to a model containing *fluency and connectedness*, the bottom portion represents the ΔR^2 for *fluency and connectedness* when added to a model containing the indicated variable, and the middle part indicates their overlap. Thus, the r^2 (the simple correlation squared for the one-predictor model) for *fluency and connectedness* is the bottom plus the middle portions; similarly, the r^2 for the indicated variable is the middle plus the top portions.

In the low-INR and high-INR groups—but not the middle-INR group—*fluency and connectedness* accounted for significant additional variance in children's 36-month expressive language, above and beyond the five variables indicated in Figure 2. In the high-INR group—but again not the middle-INR group—*fluency and connectedness* accounted for significant additional variance in children's 36-month receptive language, above and beyond the same five variables indicated in Figure 3, and in the low-INR group, only above and beyond maternal WPM and child's CDI count. In the middle-INR group, the child's CDI count, but not *fluency and connectedness*, accounted for significant additional variance for both 36-month expressive and receptive language. See Table 5 for the results of the regressions.

3.3 | Controlling for 24-month child language

The bars presented in Figures 2 and 3 represent 2-predictor models. Of the five bars for each INR group, two predict 36-month child language outcomes from 24-month child language variables, but the other three predict from 24-month maternal language variables and a parenting composite based largely on earlier measures. It is not a given that we should always control for a highly correlated precursor when predicting an outcome; it can account for so much of the variance that effects of other variables can be obscured rather than illuminated. Nonetheless, as an exploratory question we

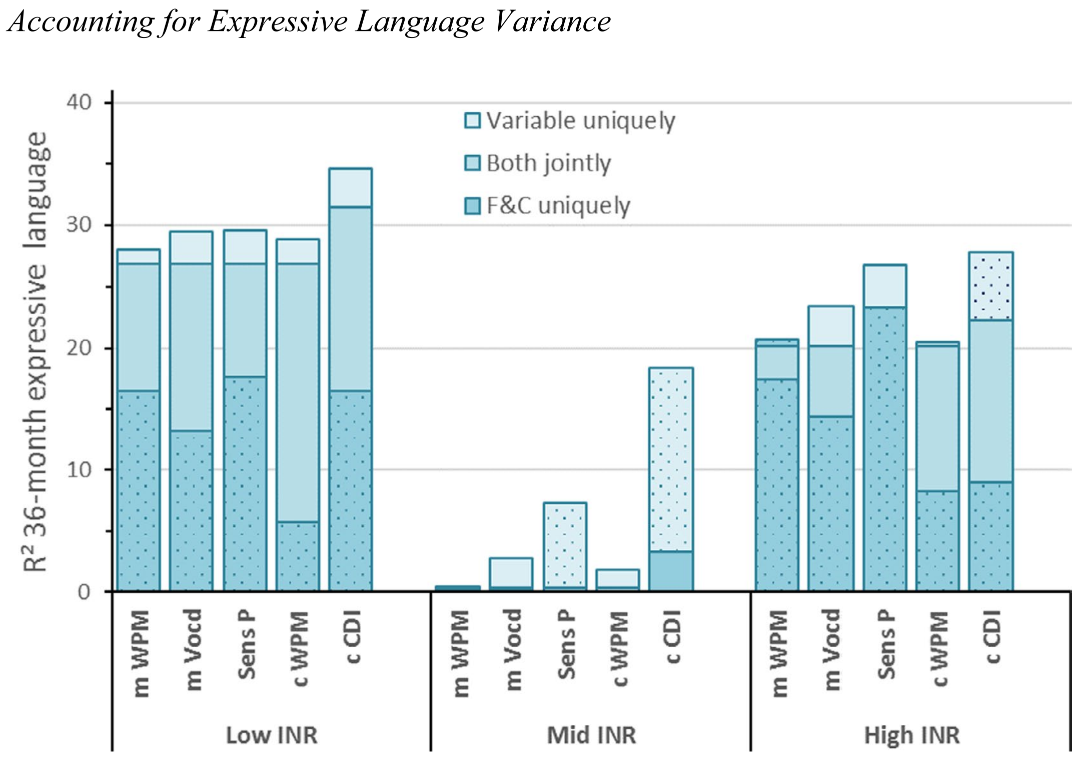

FIGURE 2 Accounting for expressive language variance. *Note*. The bars in this figure show how expressive language variance was apportioned between *fluency and connectedness* (F&C) and the specified variable within the low-, middle-, and high-INR groups. The height of each bar indicates the amount of variance accounted for by a two-predictor model (total R^2, here expressed as a percent). The predictor variables are maternal WPM (m WPM), maternal lexical diversity (m VocD), sensitive parenting (Sens P), child words per minute (c WPM), and child CDI counts (c CDI). Each bar is divided into three parts. The bottom part represents the portion of total R^2 uniquely accounted for by F&C, the top part represents the portion of total R^2 uniquely accounted for by specified variable, and the middle part indicates the portion of total R^2 accounted for by the overlap between the two variables. Thus the bottom and middle portions together represent the total amount of variance accounted for by F&C. Unique portions significant at the .05 level are dotted. See text for further details

asked what the results would be if the models predicting 36-month child language from maternal and parenting variables included—that is, controlled for—24-month child language. The answer for these 3-predictor models is that the ΔR^2 s for *fluency and connectedness* were reduced—as expected when additional partially redundant variables are added to a regression model—but most of those that were statistically significant without control remained statistically significant with control.

For 36-month expressive language in both low- and high-INR groups, the *p* values associated with the ΔR^2 s for *fluency and connectedness* when added to each of the three predictor variables (maternal WPM, maternal lexical diversity, and sensitive parenting) were all less than .05 (see Figure 2). They remained so when controlling for CDI counts in both INR groups and when controlling for child WPM in the high-INR group; in the low-INR group, *p* values were .10, .16, and .06, respectively, approaching but not reaching conventional statistical significance. For 36-month receptive language in the high-INR group, the *p* values associated with the ΔR^2 s for *fluency and connectedness* when added to each of the three predictor variables were all less than .05, as were the *p* values associated with the ΔR^2 s for *fluency and connectedness* when added to maternal WPM in the low-INR group (see Figure

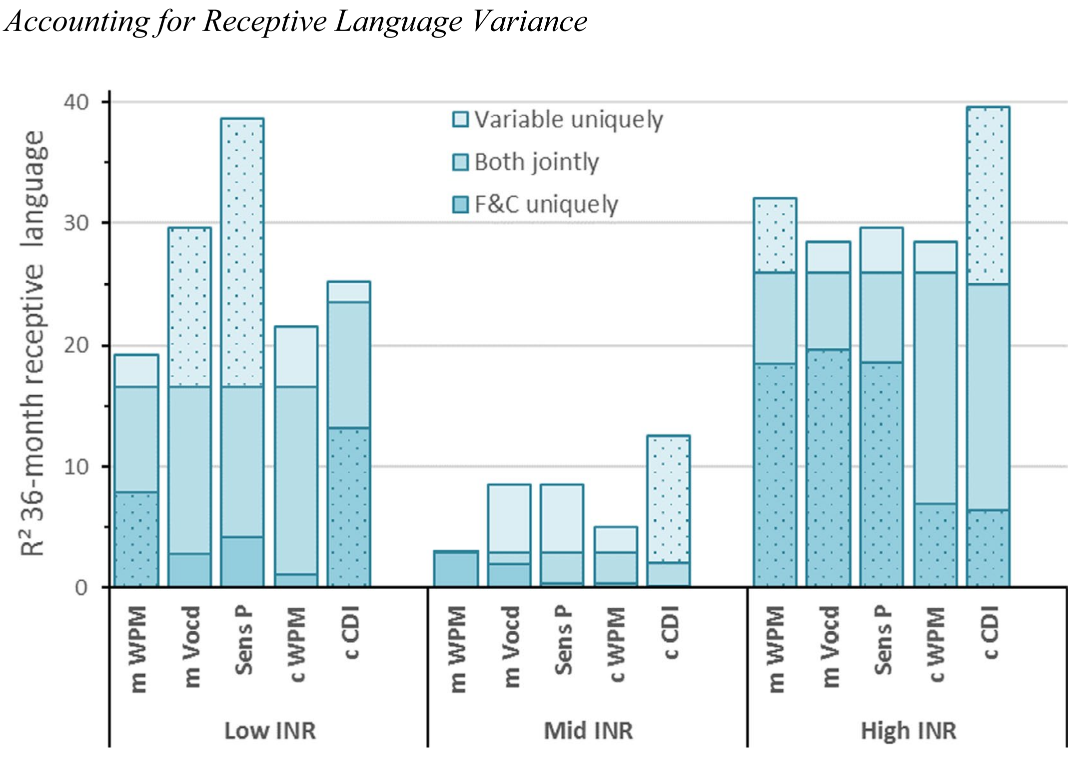

FIGURE 3 Accounting for receptive language variance. *Note*. The bars in this figure show how receptive language variance was apportioned between *fluency and connectedness* (F&C) and the specified variable within the low-, middle-, and high-INR groups. The height of each bar indicates the amount of variance accounted for by a two-predictor model (total R^2, here expressed as a percent). The predictor variables are maternal WPM (m WPM), maternal lexical diversity (m VocD), sensitive parenting (Sens P), child words per minute (c WPM), and child CDI counts (c CDI). Each bar is divided into three parts. The bottom part represents the portion of total R^2 uniquely accounted for by F&C, the top part represents the portion of total R^2 uniquely accounted for by the specified variable, and the middle part indicates the portion of total R^2 accounted for by the overlap between the two variables. Thus the bottom and middle portions together represent the total amount of variance accounted for by F&C. Unique portions significant at the .05 level are dotted. See text for further details

3). They remained so in the high-INR group (albeit $p = .061$ for sensitive parenting when controlling for child WPM) and in the low-INR group when controlling for child CDI count but not child WPM.

In sum, when predicting expressive language, the pattern of statistical significance was not affected when controlling for CDI (6 of 6 remained significant) but was reduced when controlling for child WPM (3 of 6 remained significant, although $p = .060$ for a fourth). Corresponding numbers when predicting receptive language were 4 of 4 for CDI and 2 of 4 for child WPM, although $p = .061$ for a third.

4 | DISCUSSION

In this study, we expanded upon prior research by asking whether more fluid and connected caregiver–toddler interaction at age 2 years predicted better language outcomes at age 3 years for children from families with higher household incomes than the low-income families reported by Hirsh-Pasek et al. (2015). By examining children with a range of language outcomes across income level, this research

TABLE 5 Accounting for variance in expressive and receptive language

	Low-INR group			Middle-INR group			High-INR group		
Variable	Variable R^2	F&C ΔR^2	Total R^2	Variable R^2	F&C ΔR^2	Total R^2	Variable R^2	F&C ΔR^2	Total R^2
Expressive language									
Maternal WPM	**11**	**16**	**28**	0.3	0.2	0.5	3.3	**17**	**21**
Maternal VocD	**16**	**13**	**29**	2.6	0.1	2.7	**9.1**	**14**	**23**
Sensitive parenting	**12**	**18**	**30**	**7.0**	0.3	7.3	3.4	**20**	**24**
Child WPM	**23**	**5.7**	**29**	1.8	0.1	1.8	**12**	**8.2**	**21**
CDI word count	**18**	**16**	**35**	**12**	3.3	**15**	**14**	**9.0**	**28**
Receptive language									
Maternal WPM	**11**	**7.9**	**19**	0.1	2.9	2.9	**14**	**18**	**32**
Maternal VocD	**27**	2.7	**30**	**6.6**	1.9	8.5	**8.8**	**20**	**28**
Sensitive parenting	**34**	4.1	**39**	**8.2**	0.3	8.5	**11**	**19**	**30**
Child WPM	**20**	1.1	**22**	4.7	0.3	5.0	**22**	**6.9**	**28**
Raw CDI score	**12**	**13**	**25**	**12**	0.1	**13**	**33**	**6.4**	**40**

Note: $n = 60$ for each INR group except for the CDI score, for which $n = 53$, 55, and 57 for the low-, middle-, and high-INR group, respectively. The R^2 for the specified variable, the ΔR^2 when fluency and connectedness (F&C) is added, and the total R^2 are from stepwise, hierarchic multiple regressions. Statistically significant ($p < .05$) proportions of variance are bolded. Totals may not sum exactly due to rounding. Similar statistics for maternal WPM and sensitive parenting in the low-income group were reported in Hirsh-Pasek et al. (2015).

Abbreviations: CDI, Communication Development Inventory; WPM, words per minute.

answers questions about the role of *fluency and connectedness* in relation to other input features such as quantity and quality of language input, as well as sensitive parenting. Like very few other studies (e.g., Cartmill et al., 2013), we examined variation *within* income level and asked whether the same mechanisms fed learning outcomes across income groups.

We first asked whether the finding from Hirsh-Pasek et al. (2015)—that *fluency and connectedness* at 24 months predicted language at 36 months in a low-INR sample—extended to samples of middle- and high-INR toddlers. We hypothesized that it would, but this prediction was only partially supported. In the high-INR sample, *fluency and connectedness* was a moderate to strong predictor of language ($r = .45$, .51; comparable values for the low-INR group were .52, .41). This relation, however, was not found in the middle-INR group ($r = .06$, .17), a surprising finding that we discuss in depth later.

We next looked at whether the findings persisted after controlling for the amount of parent talk, the quality of parent talk (as measured by lexical diversity), and sensitive parenting. For the low-INR and high-INR toddlers, they did, but for the middle-INR group they did not. For the high-INR sample, *fluency and connectedness* uniquely predicted both expressive and receptive language over and above the other variables. For the low-INR sample, *fluency and connectedness* was a unique predictor of expressive language over and above the other variables and a unique predictor of receptive language over and above amount of talk. The results in the low- and high-INR groups suggest that the relation between *fluency and connectedness* and language skills seen in these groups could not be fully accounted for by the experience of more talk, richer talk, or overall more sensitive parenting. This result implies that the behaviors that make up *fluency and connectedness* are specific to language learning and not merely a measure of overall higher-quality parent–toddler interaction.

Finally, we examined whether *fluency and connectedness* predicted language outcomes when we controlled for earlier child language ability at 24 months. Once again, in the low-INR and high-INR samples, it did but in the middle-INR group it did not. We looked at two indicators of productive language at 24 months, talkativeness during the interaction and parent-reported expressive vocabulary, and for the low- and high-INR groups *fluency and connectedness* predicted child expressive and receptive language skills over and above these variables, with one exception. This provides evidence that the quality of communication adds some additional information to our understanding of language development over and above the child's earlier language skills.

4.1 | Implications of findings from the low- and high-INR samples

The data from the low- and high-INR samples provide evidence that the importance of language input comes not just from quantity and quality of talk or even general sensitive parenting, but from embedding talk into a meaningful, communicative context. This is supported by research using the Human Simulation Model (Cartmill et al., 2013; Gillette et al., 1999; Trueswell et al., 2016), which suggests that to learn a new word, an infant only needs to hear it once, but the word must be heard in the right context, a context in which the infant's attention, the language input, and the referent of the input are aligned. Hearing a word many times may increase the chance it is heard in the correct context, but the number of times in and of itself is less important. If an infant is consistently engaged in high-quality interaction, the chance of hearing any given word in the ideal, aligned context is higher, making the total number of words heard, up to a certain point, irrelevant. In terms of *fluency and connectedness*, it is helpful when the parents' talk is embedded in the context of the back-and-forth interaction. If the interaction is fluid and balanced, the infant is hearing language input in a way that aligns with the dyad's shared attention and the conversation is adjusted in such a way that the infant has an opportunity to actively participate.

The importance of back-and-forth interaction also relates to the idea of parental responsivity, communication in which the caregiver responds to the infant in way that is both temporally and thematically appropriate while adjusting their speech to the infant's communication level (Reed et al., 2016; Tamis-LeMonda et al., 2014). Parental responsivity has been shown to relate to children's language learning (e.g., Goldstein et al., 2003; Landry et al., 2008; Tamis-LeMonda et al., 2001) and may set the stage for fluent and connected conversation. When an adult responds in a prompt and meaningful way to an infant's explorations, the infant stays focused for longer periods of time (Mason et al., 2018; Miller et al., 2009; Miller & Gros-Louis, 2013), which may in turn allow for more of the balanced and tightly coupled back-and forth communication characterized by *fluency and connectedness*.

The importance of high-quality communication for language development is further supported by recent findings that both activation and connectivity in Broca's area, a region of the brain strongly associated with language, are related to the number of conversational turns, but not the amount of speech, that children heard (Romeo, Leonard, et al., 2018; Romeo, Segaran, et al., 2018). Although that study looked at the amount of conversation and not the quality of the conversation per se, it suggests that language-related brain development is associated with engaging in conversation *with* children, not talking *at* them.

4.2 | Small magnitude of effects in the middle-INR sample

When we considered our combined sample—which contained equal numbers of families from the three INR groups and which had been selected from a sample of families in which the three INR groups were approximately equally represented (28%, 38%, and 34% were in the low-, middle-, and high-INR groups, respectively)—we found that associations of 24-month maternal quantity and quality of talk were weakly to moderately associated with 36-month expressive and receptive language. The rating of *fluency and connectedness* was moderately associated. Yet when we considered the three INR groups separately—most household incomes less than $60,000, between $60,000 and $120,000, and over $120,000 (in 2020 dollars), respectively—a different pattern emerged for the middle-INR group. In the middle-INR group, associations of maternal quantity of talk with expressive and receptive language were near zero and associations of *fluency and connectedness* were near zero for expressive and weak for receptive language—in contrast to moderate and strong associations in the low- and high-INR groups. Only quality of maternal talk deviated a bit from this pattern. Its associations with expressive and receptive language were weak in the middle-INR group, although moderate to strong in the low- and high-INR groups. Furthermore, when multiple predictors were examined in the regression, the predictors accounted for less variance in outcomes for the middle-INR compared to the low- and high-INR. Indeed, for the middle-INR sample, the median amount of variance accounted for was 2% for expressive language and 9% for receptive language, whereas in the low- and high-INR samples, the medians were 29% and 22% for expressive language and 25% and 30% for receptive language, respectively. The question is, what accounts for the different pattern in the middle-INR group.

We propose two possible explanations. First, it could be an unlucky draw or a result of equating our samples for expressive language, but this seems unlikely. We compared means and correlations in the selected and unselected middle-INR samples and found little difference. The standardized mean differences for variables in the SECCYD archive—child CDI word count, sensitive parenting, and expressive and receptive language—were 0.01 to 0.03. And correlations of CDI counts with expressive and receptive language were moderate in both samples, ranging from .34 to .41, although correlations for sensitive parenting were weak in the selected sample (.26, .29) but still moderate in the unselected sample (.38, .47).

Second, there could be a real difference in language learning between children from middle-INR backgrounds and children from either low- or high-INR backgrounds. A meta-analysis of 37 empirical papers looking at the associations between parental responsiveness and child language found larger effect sizes in low-SES samples and diverse-SES samples than in middle-to-high-SES samples (Madigan et al., 2019)—recognizing that these groups do not necessarily align with our low-, middle-, and high-INR groups. Of the 12 studies identified as having middle-to-higher-SES samples, three studies made a broad statement about SES (e.g., mostly middle-class; Heinicke et al., 1986), three studies presented an average for indicators of SES (e.g., median income; Karrass, & Braungart-Rieker, 2003), five studies presented both an average and a range for indicators of SES (e.g., median and range for family income in dollars and parent education; Nozadi et al., 2013), and one study had two groups based on maternal education (Bee et al., 1982). None of the samples in these studies would have fit our definition of middle-INR. It is possible that measures of language input that have been found to predict language outcomes in higher-SES samples and lower-SES samples simply do not generalize to middle-SES samples, and this has yet to be evident because middle- and higher-SES groups are so often grouped together. Even in Hart and Risley's (1995) classic study, the "working class" sample was initially identified as two separate SES groups that were then combined.

To determine whether any recent studies had looked at the relation between parent–child interaction and language development in a sample similar to our middle-INR sample, we conducted an in-depth review of the literature. Using Google Scholar, we looked at studies published between 2015 and 2020 that included the terms "parent-child interaction" or "parent-infant interaction" or "parent-toddler interaction" and either "language" or "vocabulary" and either "middle-income" or "middle-SES." These searches yielded 1,648 results, none of which examined parent–child interaction at

24 months and language outcomes at 36 months in a middle-INR sample in the United States. Thus, the middle-INR group appears to be surprisingly understudied. If it is the case that the findings from the low-, middle-, and high-INR groups do indeed represent true group differences, future research needs to investigate the relation between children's early social interactions and language development in children from middle-INR families, separately from their higher- and lower-INR peers, to determine whether there are differences in the relations between features of caregiver–child interaction and language development in this group.

4.3 | Limitations

To answer the research questions posed, participants were selected to represent a wide range of abilities at all levels of family income, oversampling high-achieving, low-INR toddlers and low-achieving, high-INR toddlers. Although this allowed us to see the full range of variability within each income group and tease apart the often-intertwined effects of income level and language proficiency, this is not the same distribution as would be seen in the general population. These effects should also be investigated in a representative sample to see whether the same amount of variation exists and if there are income group-based differences in *fluency and connectedness* during parent–toddler interaction.

SES—however variously defined—is a conglomerate of factors, including income, job prestige, and education, that speak to one's access to social and economic resources. Here, we used income-to-needs ratio to sample groups classified as high-INR, middle-INR, and low-INR. Even though these groups were selected based on household income, the average education levels of the mothers were significantly different across our income groups, suggesting that our income group distinction might reflect differences in other facets of SES. Future research should further examine how other indicators of SES, such as use of government assistance, relate to the quality of caregiver–child communication.

Although the sample was diverse in terms of income, all toddlers were North American speakers of English and the majority identified as White, although percentages varied by INR group. Research suggests that there are ethnic and racial differences in talk that have implications for children's language development (e.g., Tamis-LeMonda et al., 2012). Given the differences observed across income groups in the proportion of White children, we were unable to investigate the role of racial and ethnic differences in talk. However, a recent study found that *fluency and connectedness*, in concert with other aspects of the quality of communication, predicted language and early literacy outcomes in a low-income Mexican-American Spanish-speaking sample (Adamson et al., unpublished data). Future research should examine these features of high-quality caregiver–child interaction in other samples.

Similarly, caregiver talk also varies across cultural groups of the United States. Research suggests that cultural group differences outweigh SES differences in talk (Sperry et al., 2019b). Although our sample was drawn from different areas of the United States, participants were not selected to be representative of specific cultural groups, precluding examination of these differences. Future research should also investigate cultural group differences in the *fluency and connectedness* of interactions between caregivers and children.

4.4 | Future directions

This study looked specifically at mother–toddler interaction, though evidence suggests that even within the United States infants are often cared for by a variety of adults and older children, especially in lower-SES families (Sperry et al., 2019b). Furthermore, research suggests that if children do not have a sufficiently rich early language environment in the home, either by adults or other children such as siblings, caregivers in daycare might be able to compensate (Vernon-Feagans et al., 2013). To get a fuller picture of the effects of caregiver–child interaction and quantity of language input, research should investigate how fluid and connected interactions are established by different caregivers.

Along with different caregivers, future research should consider different activities in which caregiver and child are involved. Research has shown that parent talk (Hoff-Ginsberg, 1991; Montag et al., 2018) as well as parental responsivity (Bornstein et al., 1999) and various aspects of joint engagement (Adamson et al., 2004, 2012) vary by the activity and context. Looking at *fluency and connectedness* in a wider range of activities will help to develop a better understanding of this construct.

5 | CONCLUSION

In conclusion, this study expanded upon prior research in two ways: (a) by asking whether the mechanisms for language development that work within a group for toddlers from low-SES backgrounds hold for toddlers from higher-SES backgrounds when child language outcomes are equated across income; and (b) by examining the effects of other parenting factors and earlier child language ability relative to *fluency and connectedness* on language outcome across an economically diverse sample. We found support for our hypotheses in the high-INR group but not in the middle-INR group. These findings raise the question of whether the measure of *fluency and connectedness* captures a central feature of caregiver–child interactions that promote language and suggest maybe it does not for all income groups. Indeed, the surprising lack of effect in our middle-INR group raises the question of for whom *fluency and connectedness* builds language and if the research done on language learning has adequately represented children from middle-INR homes. For children in the low- and high-INR group, the prediction from *fluency and connectedness* was robust, suggesting that a strong foundation for communication through fluid and connected interactions may boost language development for some children. Future research should aim to understand what aspects of interaction promote language learning within and across SES groups to better understand how early caregiver–child interactions support optimal language development.

ACKNOWLEDGEMENTS

This study uses data reported in Hirsh-Pasek et al. (2015) and, like that paper, draws data and uses video records from the archived NICHD Study of Early Child Care and Youth Development. This research was supported by the Temple University Graduate Fellowship awarded to LRM and grants from the Bezos Family Foundation to KHP (270642 18110-02), William Penn Foundation (#45-15) to KHP, and the Institute of Education Sciences to KHP and RMG (R305A110284, R324A160241). The authors declare no conflicts of interest with regard to the funding source for this study.

ORCID

Lillian R. Masek https://orcid.org/0000-0003-0448-3671
Sarah J. Paterson https://orcid.org/0000-0002-5261-5109
Roger Bakeman https://orcid.org/0000-0001-6517-3111
Lauren B. Adamson https://orcid.org/0000-0002-6601-8874
Margaret Tresch Owen https://orcid.org/0000-0003-0208-3965
Amy Pace https://orcid.org/0000-0003-3018-419X
Kathy Hirsh-Pasek https://orcid.org/0000-0003-2947-4544

REFERENCES

Adamson, L. B., Bakeman, R., & Deckner, D. F. (2004). The development of symbol-infused joint engagement. *Child Development*, *75*(4), 1171–1187. https://doi.org/10.1111/j.1467-8624.2004.00732.x

Adamson, L. B., Bakeman, R., Deckner, D. F., & Nelson, P. B. (2012). Rating parent-child interactions: Joint engagement, communication dynamics, and shared topics in autism, Down syndrome, and typical development. *Journal of Autism and Developmental Disorders*, *42*, 2622–2635. https://doi.org/10.1007/s10803-012-1520-1

Adamson, L. B., Bakeman, R., & Suma, K. (2016). *The Joint Engagement Rating Inventory (JERI)*. Technical Report 25. Unpublished manuscript, Department of Psychology, Georgia State University.

Adamson, L. B., Bakeman, R., Suma, K., & Robins, D. L. (2019). An expanded view of joint attention: Skill, engagement, and language in typical development and autism. *Child Development*, *90*(1), e1–e18. https://doi.org/10.1111/cdev.12973

Bakeman, R. (2018). *KappaAcc: Deciding whether kappa is big enough by estimating observer accuracy (Technical Report 28)*. Retrieved from Georgia State University website: http://bakeman.gsucreate.org/DevLabTechReport28.pdf

Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods in the behavioral sciences*. Cambridge University Press.

Bee, H. L., Barnard, K. E., Eyres, S. J., Gray, C. A., Hammond, M. A., Spietz, A. L., Snyder, C., & Clark, B. (1982). Prediction of IQ and language skill from perinatal status, child performance, family characteristics, and mother-infant interaction. *Child Development*, *53*(5), 1134–1156. https://doi.org/10.2307/1129003

Bornstein, M. H., Tamis-LeMonda, C. S., & Haynes, M. (1999). First words in the second year. Continuity, stability, and models of concurrent and lagged correspondence in vocabulary and verbal responsiveness across age and context. *Infant Behavior and Development*, *22*(1), 67–87. https://doi.org/10.1016/S0163-6383(99)80006-X

Caldwell, B. M., & Bradley, R. H. (1984). *Administration manual: Home observation for measurement of the environment*. University of Arkansas at Little Rock.

Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(28), 11278–11283. https://doi.org/10.1073/pnas.1309518110

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Conway, L., Levickis, P., Smith, J., Mensah, F., Wake, M., & Reilly, S. (2018). Maternal communicative behaviours and interaction quality as predictors of language development: Findings from a community-based study of slow-to-talk toddlers. *International Journal of Language & Communication Disorders*, *54*(2), 339–354. https://doi.org/10.1111/1460-6984.12352

Dudley-Marling, C., & Lucas, K. (2009). Pathologizing the language and culture of poor children. *Language Arts*, *86*(5), 362–370.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., & Reilly, J. S. (1991). *Technical manual for the MacArthur communicative development inventories*. San Diego State University.

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Kimbrough Oller, D., Hansen, J. H. L., & Paul, T. D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology*, *26*(2), 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176. https://doi.org/10.1016/S0010-0277(99)00036-0

Ginsborg, J. (2006). The effects of socio-economic status on children's language acquisition and use. In J. Clegg, & J. Ginsborg (Eds.), *Language and social disadvantage: Theory into practice*, (pp. 9–27). Wiley.

Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(13), 8030–8035. https://doi.org/10.1073/pnas.1332441100

Golinkoff, R. M., Hoff, E., Rowe, M., Tamis-LeMonda, C., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. *Child Development*, *90*(3), 985–992. https://doi.org/10.1111/cdev.13128

Hart, B., & Risley, R. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Paul H. Brooks.

Heinicke, C. M., Diskin, S. D., Ramsey-Klee, D. M., & Oates, D. S. (1986). Pre- and postbirth antecedents of 2-year-old attention, capacity for relationships, and verbal expressiveness. *Developmental Psychology*, *22*(6), 777–787. https://doi.org/10.1037/0012-1649.22.6.777

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., Yust, P. K. S., & Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, *26*(7), 1071–1083. https://doi.org/10.1177/0956797615581493

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368–1378. https://doi.org/10.1111/1467-8624.00612

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*(1), 4–14. https://doi.org/10.1037/a0027238

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child Development*, *62*(4), 782–796. https://doi.org/10.1111/j.1467-8624.1991.tb01569.x

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365. https://doi.org/10.1016/j.cogpsych.2010.08.002

Karrass, J., & Braungart-Rieker, J. M. (2003). Parenting and temperament as interacting agents in early language development. *Parenting: Science and Practice*, *3*(3), 235–259. https://doi.org/10.1207/S15327922PAR0303_03

Landry, S. H., Smith, K. E., Swank, P. R., & Guttentag, C. (2008). A responsive parenting intervention: the optimal timing across early childhood for impacting maternal behaviors and child outcomes. *Developmental Psychology*, *44*(5), 1335–1353. https://doi.org/10.1037/a0013030

Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *International Journal of Language and Communication Disorders*, *35*(2), 165–188. https://doi.org/10.1080/136828200247133

Leigh, P., Nievar, M. A., & Nathans, L. (2011). Maternal sensitivity and language in early childhood: A test of the transactional model. *Perceptual and Motor Skills*, *113*(1), 281–299. https://doi.org/10.2466/10.17.21.28.PMS.113.4.281-299

Levine, D., Pace, A., Luo, R., Hirsh-Pasek, K., Golinkoff, R. M., de Villiers, J., Iglesias, A., & Wilson, M. S. (2020). Evaluating socioeconomic gaps in preschoolers' vocabulary, syntax, and language process skills with the Quick Interactive Language Screener (QUILS). *Early Childhood Research Quarterly*, *50*, 114–128. https://doi.org/10.1016/j.ecresq.2018.11.006

Luo, R., Pace, A., Masek, L., Hirsh-Pasek, K., & Golinkoff, M. R. (2016). The family's role in the relation between socioeconomic status and early language development. *Journal of Family Medicine*, *3*(6), 1073–1077.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Lawrence Erlbaum.

Madigan, S., Prime, H., Graham, S. A., Rodrigues, M., Anderson, N., Khoury, J., & Jenkins, J. M. (2019). Parenting behavior and child language: A meta-analysis. *Pediatrics*, *144*(4), e20183556. https://doi.org/10.1542/peds.2018-3556

Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development*. Palgrave Macmillan.

Mason, G. M., Kirkpatrick, F., Schwade, J. A., & Goldstein, M. H. (2018). The role of dyadic coordination in organizing visual attention in 5-month-old infants. *Infancy*, *24*(2), 162–186. https://doi.org/10.1111/infa.12255

Michaels, S. (2013). Commentary: Déjà vu all over again: What's wrong with Hart & Risley and a "linguistic deficit" framework in early childhood education? *Learning Landscapes*, *7*(1), 23–41.

Miller, J. L., Ables, E. M., King, A. P., & West, M. J. (2009). Different patterns of contingent stimulation differentially affect attention span in prelinguistic infants. *Infant Behavior and Development*, *32*(3), 254–261. https://doi.org/10.1016/j.infbeh.2009.02.003

Miller, J. L., & Gros-Louis, J. (2013). Socially guided attention influences infants' communicative behavior. *Infant Behavior and Development*, *36*(4), 627–634. https://doi.org/10.1016/j.infbeh.2013.06.010

Miller, J., & Iglesias, A. (2010). *Systematic analysis of language transcripts (SALT) research version*. SALT Software.

Mistry, R. S., Biesanz, J. C., Chien, N., Howes, C., & Benner, A. D. (2008). Socioeconomic status, parental investments, and the cognitive and behavioral outcomes of low-income children from immigrant and native households. *Early Childhood Research Quarterly*, *23*(2), 193–212. https://doi.org/10.1016/j.ecresq.2008.01.002

Montag, J. L., Jones, M. N., & Smith, L. B. (2018). Quantity and diversity: Simulating early word learning environments. *Cognitive Science*, *42*, 375–412. https://doi.org/10.1111/cogs.12592

National Institute of Child Health and Human Development Early Child Care Research Network. (1999). Child care and mother-child interaction in the first three years of life. *Developmental Psychology*, *35*(6), 1399–1413. https://doi.org/10.1037/0012-1649.35.6.1399

National Institute of Child Health and Human Development Early Child Care Research Network. (2002). Early child care and children's development prior to school entry: Results from the NICHD Study of Early Child Care. *American Educational Research Journal*, *39*(1), 133–164. https://doi.org/10.3102/00028312039001133

National Institute of Child Health and Human Development Early Child Care Research Network. (2006). Child-care effect sizes for the NICHD Study of Early Child Care and Youth Development. *American Psychologist*, *61*(2), 99–116. https://doi.org/10.1037/0003-066X.61.2.99

Nozadi, S. S., Spinrad, T. L., Eisenberg, N., Bolnick, R., Eggum-Wilkens, N. D., Smith, C. L., Gaertner, B., Kupfer, A., & Sallquist, J. (2013). Prediction of toddlers' expressive language from maternal sensitivity and toddlers' anger expressions: A developmental perspective. *Infant Behavior and Development*, *36*(4), 650–661. https://doi.org/10.1016/j.infbeh.2013.06.002

Pace, A., Alper, R., Burchinal, M. R., Golinkoff, R. M., & Hirsh-Pasek, K. (2019). Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. *Early Childhood Research Quarterly*, *46*(1), 112–125. https://doi.org/10.1016/j.ecresq.2018.04.001

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, *3*, 285–308. https://doi.org/10.1146/annurev-linguistics-011516-034226

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. https://doi.org/10.1111/1467-8624.00498-i1

Purpura, D. J. (2019). Language clearly matters; methods matter too. *Child Development*, *90*(6), 1839–1846. https://doi.org/10.1111/cdev.13327

Raviv, T., Kessenich, M., & Morrison, F. J. (2004). A mediational model of the association between socioeconomic status and three-year-old language abilities: The role of parenting factors. *Early Childhood Research Quarterly*, *19*(4), 528–547. https://doi.org/10.1016/j.ecresq.2004.10.007

Reed, J., Hirsh-Pasek, K., & Golinkoff, R. M. (2016).Meeting children where they are: Adaptive contingency builds early communication skills. In P. Witt (Ed.) *Communication and learning (Vol. 16, Handbooks of communication science)* (pp. 601–627). : deGruyter Mouton.

Reynell, J. (1991). *Reynell developmental language scales (U.S. Edition)*. : Western Psychological Services.

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700–710. https://doi.org/10.1177/0956797617742725

Romeo, R. R., Segaran, J., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Yendiki, A., Rowe, M. L., & Gabrieli, J. D. E. (2018). Language exposure relates to structural neural connectivity in childhood. *Journal of Neuroscience*, *38*(36), 7870–7877. https://doi.org/10.1523/JNEUROSCI.0484-18.2018

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*(5), 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Smith, J., Eadie, T., Levickis, P., Bretherton, L., & Goldfeld, S. (2018). Predictive validity of verbal and non-verbal communication and mother–child turn-taking at 12 months on language outcomes at 24 and 36 months in a cohort of infants experiencing adversity: A preliminary study. *International Journal of Language & Communication Disorders*, *53*(5), 969–980. https://doi.org/10.1111/1460-6984.12408

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019a). Language does matter: But there is more to language than vocabulary and directed speech. *Child Development*, *90*(3), 993–997. https://doi.org/10.1111/cdev.13125

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019b). Reexamining the verbal environments of children from different socioeconomic backgrounds. *Child Development*, *90*(4), 1303–1318. https://doi.org/10.1111/cdev.13072

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development*, *72*(3), 748–767. https://doi.org/10.1111/1467-8624.00313

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science*, *23*(2), 121–126. https://doi.org/10.1177/0963721414522813

Tamis-LeMonda, C. S., Song, L., Leavell, A. S., Kahana-Kalman, R., & Yoshikawa, H. (2012). Ethnic differences in mother–infant language and gestural communications are associated with specific skills in infants. *Developmental Science*, *15*(3), 384–397. https://doi.org/10.1111/j.1467-7687.2012.01136.x

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*(6), 1454–1463. https://doi.org/10.2307/1130423

Trueswell, J. C., Lin, Y., Armstrong, B., Cartmill, E. A., Goldin-Meadow, S., & Gleitman, L. R. (2016). Perceiving referential intent: Dynamics of reference in natural parent–child interactions. *Cognition*, *148*, 117–135. https://doi.org/10.1016/j.cognition.2015.11.002

Vanormelingen, L., & Gillis, S. (2016). The influence of socio-economic status on mothers' volubility and responsiveness in a monolingual Dutch-speaking sample. *First Language*, *36*(2), 140–156. https://doi.org/10.1177/0142723716639502

Vernon-Feagans, L., Bratsch-Hines, M. E., &The Family Life Project Key Investigators. (2013). Caregiver–child verbal interactions in child care: A buffer against poor language outcomes when maternal language input is less. *Early Childhood Research Quarterly*, *28*(4), 858–873. https://doi.org/10.1016/j.ecresq.2013.08.002

How to cite this article: Masek LR, Paterson SJ, Golinkoff RM, et al. Beyond talk: Contributions of quantity and quality of communication to language success across socioeconomic strata. *Infancy*. 2020;26:123–147. https://doi.org/10.1111/infa.12378