Journal of Child Language (2022), **49**, 1037–1051
doi:10.1017/S0305000921000441

CAMBRIDGE
UNIVERSITY PRESS

BRIEF RESEARCH REPORT

Socioeconomic status correlates with measures of Language Environment Analysis (LENA) system: a meta-analysis

Leonardo PIOT[1], Naomi HAVRON[1,2], and Alejandrina CRISTIA[1]

[1]Laboratoire de Sciences Cognitives et de Psycholinguistique, Département d'Etudes cognitives, ENS, EHESS, CNRS, PSL University, France and [2]University of Haifa, Israel
Address for correspondence: Leonardo Piot, 29 rue d'Ulm, 75005, Paris, France.
E-mail: leo1997p@gmail.com

(Received 30 November 2020; revised 14 April 2021; accepted 15 May 2021;
first published online 25 June 2021)

Abstract

Using a meta-analytic approach, we evaluate the association between socioeconomic status (SES) and children's experiences measured with the Language Environment Analysis (LENA) system. Our final analysis included 22 independent samples, representing data from 1583 children. A model controlling for LENA™ measures, age and publication type revealed an effect size of r_z= .186, indicating a small effect of SES on children's language experiences. The type of LENA metric measured emerged as a significant moderator, indicating stronger effects for adult word counts than child vocalization counts. These results provide important evidence for the strength of association between SES and children's everyday language experiences as measured with an unobtrusive recording analyzed automatically in a standardized fashion.

Keywords: SES; LENA; experiences; production; input; language environment; meta-analysis

Introduction

While all typically-developing children quickly become competent users of their language, important individual and group variation is observed in the pace and trajectory of language development (e.g., Caselli *et al.*, 1995). Factors accounting for the variance in early language development include the quantity and quality of linguistic input that young children received in their home environments (Weisleder & Fernald, 2013), which may correlate with socioeconomic status (Golinkoff, Hoff, Rowe, Tamis-LeMonda & Hirsh-Pasek, 2019). In this paper, we use a meta-analytic methodology to look at socioeconomic status (SES) and how it correlates with quantitative measures of language experiences drawn using an innovative technology consisting of automatized analyses of daylong recordings (LENA™).

In the following introduction, we first discuss the ongoing debate on the magnitude (or existence) of SES differences in children's language experiences. Then, we present the Language Environment Analysis (LENA^{TM}) technology and its three main metrics of language experiences: adult word counts, conversational turn counts and child vocalization counts. Finally, we integrate these two sections to formulate our main research questions.

SES and language environment

Socioeconomic status is a broad and complex construct reflecting the social and economic resources of a household. In child research, this construct is often captured using single measures, such as parental education, income, or composites derived from both of these factors and others (e.g., parental occupation). Although input quantity varies within a given SES group (Sperry, Sperry & Miller, 2019; Weisleder & Fernald, 2013), there is some evidence suggesting that SES correlates with input quantity. In a landmark study examining vocabulary development, Hart and Risley (1995, henceforth H&R) assessed the number of words spoken to young children from socio-economically diverse families. They calculated the mean number of words spoken to each child in their study across one-hour monthly observations from the child's first to third birthday. Based on these data, they estimated that, by the age of four years, children having grown up in a high socio-economic status context will have heard three times more child-directed speech than children in a lower socio-economic status context (about 45 versus 15 million words, respectively). This finding, commonly referred to as the "30-million-word gap", was bolstered by subsequent studies which also found SES differences in the quantity and quality of child-directed speech (e.g., Hoff, 2003; Huttenlocher, Waterfall, Vasilyeva, Vevea & Hedges, 2010). Moreover, SES was found to affect language development primarily via the child's language environment (Hoff, 2003).

These three studies were all longitudinal and assessed language environment through observations. These studies' methods have been criticized. For example, in the H&R study, researchers avoided recording family interactions not involving the child, included only the language addressed to the child by one unique parent in their counts, and instructed their observers to interact as little as possible with the family during observations. The presence of an observer may have influenced caregivers' behaviors and limited the ecological validity of their data (Dudley-Marling & Lucas, 2009). This is because the presence of observers has been found to significantly and positively affect upper-middle-class mothers' interactions with their children, leading to more frequent interactions (Zegiob, Arnold & Forehand, 1975), while it is not obvious that it would have the same effect on lower-SES families. Furthermore, the decision to only consider one parent's child-directed words may have failed to account for common structural differences between socioeconomically diverse groups. Different SES groups might exhibit differences in the number and identity of people living with the child (Glick, 1976), and how frequently they address the child, in addition to a different quantity of overheard speech. These and other considerations led Sperry *et al.* (2019) to revisit H&R's conclusions. They included H&R's data but also additional recordings from five communities, two described as "poor", two "working class", and one "middle class". Contrary to H&R, observers were encouraged to interact with the family. The authors assessed children's language input in three ways: (1) speech by the primary caregiver to the

https://doi.org/10.1017/S0305000921000441 Published online by Cambridge University Press

child (H&R's definition), (2) all speech directed to the child, and (3) overheard speech in addition to speech directed to the child. They only found significant statistical differences between two of the H&R groups when using H&R's own definition, but not among any of the other groups. For the second and third definitions, the authors do not report statistical comparison tests. To try to quantify this more precisely, we calculated the correlation between a four-level SES split encompassing SES levels from both these studies (1 = "poor"/"welfare", 2 = "working class", 3 = "middle class", 4 = "professional") and the number of words in the three definitions. The Pearson correlation coefficient was $r = .608$ for the first definition, which includes H&R's four groups and Sperry *et al.*'s five groups, whereas the other definitions only include the latter, and their coefficients are $r = -.185$ and $r = .182$ respectively.

However, Golinkoff *et al.* (2019) criticized Sperry and colleagues' study for the absence of a high SES group (i.e., corresponding to the "professional" group in H&R). Furthermore, the decision to direct observers "to interact with adults and children in an interested and relaxed manner" (Sperry *et al.*, 2019; p. 1310), perhaps resulted in unrepresentative observed language input of participants' everyday experience. As a result, there is an ongoing debate in the literature regarding the association between SES and children's language experiences.

LENA Technology

One important recent technological advance in the psycholinguistic field has been the development of the Language Environment Analysis (LENA™) device and associated software. The device is a small digital recorder, worn by the child in a specially designed vest, which records audio for up to 16 hours (or even 24, in the most recent hardware). The associated software generates an automatic analysis of the amount of language occurring in the child's environment, yielding three key estimates: the number of words spoken by any adult in the near presence of the child (Adult Word Count, AWC); the number of times a child made any kind of linguistically relevant vocalization (i.e., speech or babble, but excluding vegetative noises and crying; Child Vocalization Count, CVC); and the number of times there was an adult vocalization within five seconds of a child vocalization (Conversational Turn Count, CTC). Notice that this entails a shift in how language experiences are measured, compared to other methodologies relying on human annotations of recordings: AWC likely represents overall input; CTC represents back-and-forth interactions between the key child and adults in the environment, and thus closer to child-directed speech; and CVC captures the child's own productions, which are both necessary for establishing CTC and contribute to children's experience of their own productions. Thus, the Language ENvironment Analysis also considers this aspect of the child's experiences.

LENA™ has been found to be fairly accurate and reliable. In a recent meta-analysis comparing LENA™'s automated measures to human transcriptions, Cristia, Bulgarelli, and Bergelson (2020) found a low-to-moderate correlation for CTC ($r = .36$) but high mean correlations coefficients for AWC ($r = .79$) and CVC ($r = .77$). Moreover, the LENA™ metrics have been meta-analytically shown to predict concurrent and later language outcomes (Wang, Williams, Dilley & Houston, 2020). This meta-analysis found a medium significant association across the three measures, with the association between language outcomes and CTC/CVC (Pearson's $r = .31$ and .32 respectively) being larger than the one for AWC ($r = .21$).

https://doi.org/10.1017/S0305000921000441 Published online by Cambridge University Press

Importantly for our purposes, LENA technology can meaningfully contribute to the SES debate by providing data in which observer effects may be minimized, since there is no human observer and the sheer length of the audiorecording may contribute to habituation by the family. Moreover, research in this technique is standardized (the same hardware and software are used by all researchers), allowing an apples-to-apples integration of the previous body of data. Lastly, LENA™ is widely and increasingly used, including in intervention work (e.g., Leung, Hernandez & Suskind, 2020). As a result, our findings stand to be informative for this growing body of work.

The present study

The recent debate between Golinkoff *et al.* (2019) and Sperry *et al.* (2019) sparked an important discussion regarding the empirical basis and associated assumptions that underlie key findings in child development, and notably the purported association between SES and language experiences. Accurately measuring this association has important theoretical and practical implications, a point on which Golinkoff *et al.* (2019) and Sperry *et al.* (2019) agree. The present study therefore aims to assess to what extent SES correlates with variability in children's language experiences using LENA™ estimates, which we hope will more faithfully represent children's everyday experiences. Our questions and hypotheses were:

- What characterizes studies that report on the association between SES and LENA™'s measures? We describe publication status, country of data collection, and SES range, among other key characteristics. We are especially interested in the breadth of coverage of the data.
- What is the size of the correlation between SES and LENA™ measures?
- Are there factors leading to variation in the size of this correlation? We investigate the type of LENA™ measures (AWC, CVC, CTC), and child age, among other characteristics, as potential moderators. Given Golinkoff *et al.* (2019)'s analyses showing that SES effects are strongest when considering child-directed speech, we expect CTC to be most sensitive to SES (since it is often described as being speech engaging the child) and to CVC the least (since it captures the child's own productions). Age could be a moderator because some research suggests that the link between SES and lexical processing increases with age (Scaff, 2019).

Methods

We follow the PRISMA guidelines for systematic reviews and meta-analyses (Moher, Liberati, Tetzlaff, Altman & The PRISMA Group, 2009). All materials (including data and scripts) are available from the Open Science Framework (Piot, Havron & Cristia, 2020).

Search strategy

Searches were carried out between March 16th and July 15th, 2020, and include 1) a search on PubMedCentral with the keywords "LENA" + "socio-economic" + "language", resulting in 125 hits; 2) all the studies reported in the LENA research publications bibliography on the LENA foundation website, resulting in 105 papers; 3) a search in Google scholar using the private navigation mode (for reproducibility purposes) with the keywords "LENA", "CTC" and "socioeconomic"; 4) another search in Google

Scholar with the keywords "socioeconomic", "LENA" and ("AWC" or "CTC" or "CVC"). For both Google searches, the first hundred hits were included for screening. Finally, ten additional studies were screened based on expert advice.

Eligibility criteria

Included studies reported an association between SES and at least one of the following: AWC, CTC, and CVC. Studies where the automated measures had been manipulated (e.g., splitting AWC into child-directed speech versus overheard speech) were excluded. To include studies using daylong recordings during naturalistic interactions at home between caregivers and typically-developing children below 18 years of age, we excluded post-treatment measures from intervention studies; and atypical samples. When multiple studies used the same data (e.g., the normative LENA™ database – Gilkerson & Richards, 2008), the study with the biggest sample size and the most detailed data was included. Studies in which SES groups were confounded with another variable (e.g., country differences) were excluded. Finally, studies based on a homogeneous SES sample, in which more than 80% of the participants belonged to the same SES group, were excluded.

Study selection

See Figure 1 for PRISMA flowchart. The first author carried out all data screening and inclusion. Screening was based on title, abstract and a control + F search in the pdf for "SES", "education" and "income". This text-search method was used because many papers did not mention SES in their abstract but nonetheless referred to SES or its proxies in the paper. Twenty percent of papers were randomly selected from the first search strategy hits and screened by the second author, revealing a 96.8% screening accuracy.

Data collection process and coding

If a study reported more than one measure (i.e., AWC, CVC, and CTC), all were coded, similarly to if they reported multiple measures of SES (e.g., education and income). We extracted:

- Paper characteristics: Authors, year of publication, publication type, country, total sample size (N), subgroups and sample size/subgroups
- Method/participant characteristics: type and measure of the dependent variable used, type of SES measure (education, income, a composite) and range of SES, age, native language, gender and ethnicity of participants
- Findings: AWC, CTC, CVC quantity measures and/or correlation/effect sizes

For all the included studies, information about publication, method, and findings were extracted and coded by the first author. We then hired a freelancer to independently enter all data. There were 21 discrepancies out of 405 fields entered, mostly in descriptive fields. The first author verified all discrepancies against the original papers and corrected them if necessary.

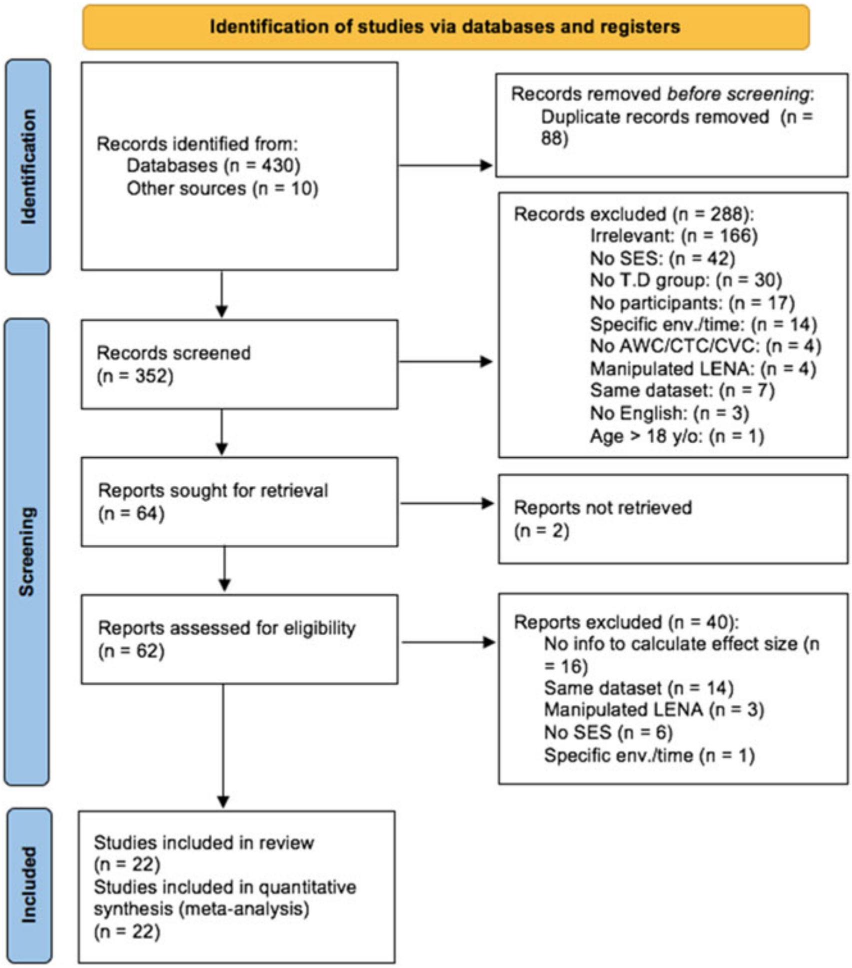

Figure 1. PRISMA Flowchart indicating studies discovered, screened, and included. "No SES" indicates studies that do not report any measures of SES and/or SES too homogeneous (>80% of participants in the same SES group) and/or SES was confounded with another variable. "No T.D. group" indicates there was no group data for typically-developing children specifically. "No participants" indicates that no participants were recruited for the study. "Specific env./time" indicates that LENA devices were only used during specific times of the day (e.g., meal-time) or specific environments (e.g., NICU). "Manipulated LENA" indicates that LENA metrics were transformed (e.g., AWC was divided in child-directed and adult-directed speech). "Same dataset" indicates that the study's dataset was the same as in another study. "No English" indicates that the paper was not written in English.

Analysis plan

All analyses were conducted with the R (Version 3.6.3; R Core Team, 2020) metafor package (Viechtbauer, 2010). We first converted raw data into Pearson's correlations (Pearson's r) using raw input quantity metrics and their variance when reported. If a study reported metrics for more than two SES groups, only the data from the lowest and highest SES groups is used to calculate effect size. This only occurred in two papers, and analyses excluding these led to the same conclusions (see Supplementary Materials, Piot, Havron & Cristia, 2020). If a study reported a correlation (Pearson's r or Spearman's rho) and not the raw data, we used the reported correlation. We treated Spearman's rho as an imperfect estimate of Pearson's r. Analyses were performed using z-transformed correlation coefficients (henceforth r_z). Positive effect sizes indicate that an increase in SES is associated with an increase in automated measures.

All the main analyses were performed using multilevel models, which account for interdependent effect sizes (e.g., AWC and CVC, education and income). Because one cannot perform a regression test for asymmetry in such a dataset, we assessed the risk for publication bias using a subset of data corresponding to AWC measures.

Results

Full output of all analyses can be accessed in the online Supplementary Materials (Piot *et al.*, 2020). We included a total of 22 studies. We classified as "published" nine journal publications that provided enough information to have an effect size (Brushe, Lynch, Reilly, Melhuish & Brinkman, 2020; d'Apice, Latham & Stumm, 2019; Dwyer, Jones, Davis, Kitamura & Ching, 2019; Ferjan-Ramírez, Lytle & Kuhl, 2020; Gilkerson, Richards, Warren, Montgomery, Greenwood, Kimbrough Oller, Hansen & Paul, 2017; Leung *et al.*, 2020; McGillion, Pine, Herbert & Matthews, 2017; Merz, Maskus, Melvin, He & Noble, 2020; Sultana, Wong & Purdy, 2020). We classified as unpublished the rest of the studies (13), because the association between SES and LENA measures was not peer-reviewed: eight cases in which the authors provided us with effect sizes, including six journal publications (Adams *et al.*, 2018; Beecher & Van Pay, 2019; Christakis, Lowry, Goldberg, Violette & Garrison, 2019; Ganek, Smyth, Nixon & Eriks-Brophy, 2018; Orena, Byers-Heinlein & Polka, 2019; Swanson, Donovan, Paterson, Wolff, Parish-Morris, Meera, Watson, Estes, Marrus, Elison, Shen, McNeilly, MacIntyre, Zwaigenbaum, St John, Botteron, Dager, Piven & IBIS Network, 2019), one SCRD poster (Romeo, Leonard, Mackey, Rowe & Gabrieli, 2019) and one evaluation report (Law, Charlton & Rush, 2018); as well as correlations extracted from two theses (PhD thesis: Lease-Johnson, 2018; MSc thesis: Flood, 2015), and from our own re-analyses of archived data (Bergelson Seedlings: Bergelson, 2017; Cougar: VanDam, 2018; Warlaumont: Warlaumont, Pretzer, Walle, Mendoza & Lopez, 2016).

Table 1 shows characteristics of the included studies. Almost all of the studies took place in an English-speaking country, with 14 studies conducted in the USA, three in the United Kingdom, two in Australia, one in Canada, one in New Zealand, and one in Vietnam. Samples were varied in terms of age range, SES, and ethnicity. Infants' mean age ranged from 5.81 to 84.36 months, although the majority (90.91% of the studies) were under 42 months. SES ranges are large, going from less than a high school diploma or 10,000US$/year of income to doctoral degrees or over 100,000US $/year. Studies tended to have a medium to large sample size, with a mean of 71.95 children (range = 12–245, total = 1583).

Turning to our quantitative synthesis, we fit a model with the following factors: mean age of children (centered) in interaction with LENA's measure (AWC as baseline, CTC, CVC), and publication type (published as baseline, unpublished). The test for moderators was significant, QM $(df = 6) = 12.95$, $p = .044$, suggesting that moderators explained variance. However, there is still variance unaccounted for in

https://doi.org/10.1017/S0305000921000441 Published online by Cambridge University Press

Table 1. Included studies. Studies identified by first author name and year regardless of the number of authors, except for the three corpora from HomeBank, identified by the corpus name. "Sample" indicates characteristics of the infant sample. "LENA" indicates characteristics of the LENA recordings ("est." stands for estimate, "h" for hourly, "12h" for 12h projection, "raw 16h" for the raw 16h counts, "z12h" for the z-scored 12h projection, "pcl" for percentiles). SES indicates characteristics of the SES measures (range); "S.D." indicates socioeconomic deprivation, "HI" Hollingshead Index, "composite" a composite measure created by the researchers.

Study	Sample	LENA	SES	Measures
Adams 2017	N = 56 16 mo (USA)	h est. (1 rec.)	HI [22 ; 66] mat. ed. [12 ; 18]	AWC CTC CVC
Adams 2017	N = 56 18 mo (USA)	h est. (1 rec.)	HI [22 ; 66] mat. ed. [12 ; 18]	AWC CTC CVC
Beecher 2019	N = 105 15 mo (USA)	raw 16h (1 rec.)	income [<31860 ; >113628]	AWC CTC
Bergelson	N = 44 6.5 mo (USA)	h est. (2 rec.)	mat. ed. [12 ; 20]	AWC CTC CVC
Brushe 2020	N = 245 5.8 mo (Australia)	raw 16h (1 rec.)	mat. ed. [HS- ; BA+]	AWC CTC CVC
Brushe 2020	N = 245 12 mo (Australia)	raw 16h (1 rec.)	mat. ed. [HS- ; BA+]	AWC CTC CVC
Christakis 2019	N = 61 5.9 mo (USA)	z12 h est. (2 rec.)	income [<10000 ; >100000] par. ed. [HS ; PG+]	AWC CTC CVC
Cougar	N = 34 26.4 mo (USA)	h est. (7 rec.)	mat. ed. [8 ; 20]	AWC CTC CVC
D'Apice 2019	N = 101 33.2 mo (UK)	12h est. (3 rec.)	composite [−1.69 ; 0.95]	AWC

(*Continued*)

Table 1. (*Continued.*)

Study	Sample	LENA	SES	Measures
Dwyer 2019	N = 50 7.9 mo (Australia)	12h est. (2 rec.)	mat. ed. [HS ; BA]	AWC CTC CVC
Dwyer 2019	N = 50 14.5 mo (Australia)	12h est. (2 rec.)	mat. ed. [HS ; BA]	AWC CTC CVC
Ferjan-Ramirez 2020	N = 71 6 mo (USA)	12h est. (2 rec.)	HI [30 ; 66]	AWC CTC CVC
Flood 2015	N = 12 33.5 mo (USA)	h est. (1 rec.)	mat. ed. [HS ; BA+]	AWC CTC
Ganek 2018	N = 17 40.2 mo (Vietnam)	12h est. (3 rec.)	mat. ed. [HS- ; HS+]	CTC
Gilkerson 2017	N = 129 22.5 mo (USA)	12h est. (10 rec.)	mat. ed. [HS- ; BA]	AWC CTC CVC
Law 2018	N = 21 19.1 mo (UK)	12h est. (1 rec.)	mat. ed. [HS- ; MA]	AWC CTC CVC
Lease-Johnson 2018	N = 113 23 mo (USA)	pcl. (1 rec.)	mat. ed. [HS ; HS+]	AWC CTC
Leung 2020	N = 157 14.2 mo (USA)	12h est. (3 rec.)	mat. ed. [HS- ; BA] par. ed. [HS- ; BA]	AWC CTC
McGillion 2017	N = 140 11.1 mo (UK)	h est. (2 rec.)	composite	AWC CVC

(*Continued*)

https://doi.org/10.1017/S0305000921000441 Published online by Cambridge University Press

Table 1. (*Continued.*)

Study	Sample	LENA	SES	Measures
Merz 2020	N = 76 84.4 mo (USA)	h est. (2 rec.)	HI [6.5 ; 20] income [2880 ; 350000] par. ed. [6.5 ; 20]	AWC CTC CVC
Orena 2019	N = 21 10 mo (Canada)	raw 16h (3 rec.)	HI [31 ; 66]	AWC CTC CVC
Romeo 2018	N = 58 69.5 mo (USA)	peak h est. (2 rec.)	income [6000 ; 250000] par. ed. [HS ; PHD]	AWC CTC CVC
Sultana 2020	N = 20 39.9 mo (NZ)	h est. (2 rec.)	mat. ed. [3 ; 10] S.D. [10 ; 1]	AWC CTC
Swanson 2019	N = 29 9.9 mo (USA)	raw 16h (2 rec.)	mat. ed. [HS ; MA]	AWC CTC
Swanson 2019	N = 29 15.7 mo (USA)	raw 16h (2 rec.)	mat. ed. [HS ; MA]	AWC CTC
Warlaumont	N = 23 9.5 mo (USA)	h est. (3 rec.)	mat. ed. [12 ; 20]	AWC CTC CVC

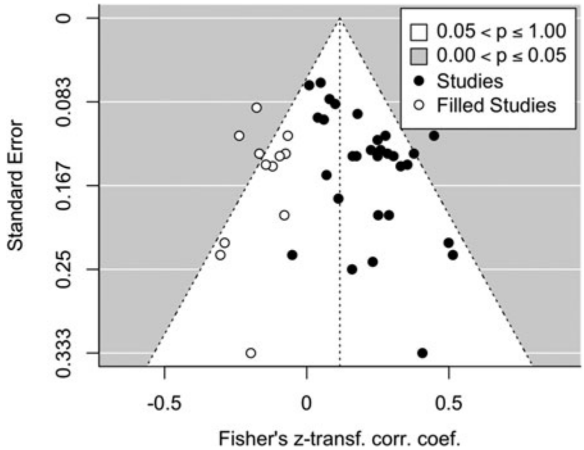

Figure 2. Funnel plot showing, for each study, its standard error as a function of its effect size. True studies are in black; those imputed by the trim-and-fill method are shown in white.

this model, since the test for heterogeneity was also significant (I2 = 40.15%, QE (df = 76) = 98.48, p = .043). The model's intercept (corresponding to the midpoint of our age distribution, the AWC measure, and published data) was significant (r_z = .186, SE = .043, p < .001, CI [.101 ; .270]), indicating a significant effect of SES on LENA measures. There was a main effect of CVC, indicating that this correlation was weaker for CVC than for the baseline AWC (r_z = −.091, SE = .034, p = .007, CI [−.158 ; −.025]). The main effect of age was marginal (r_z = .003, SE = .001, p = .086, CI [.000 ; .005]). Neither the other moderators (publication type and CTC) nor the interaction of age with measure type were significant.[1]

Publication bias

Publication bias is a risk to the validity of a meta-analysis, and thus high quality meta-analyses need to check for it, following PRISMA recommendations (Moher *et al.*, 2009). It happens when the results of a study influence the decision to publish it. If there is a publication bias in this literature favoring positive over null or negative results (Rothstein, Sutton & Borenstein, 2005), one expects studies to be asymmetrically distributed around the weighted average effect size, with high positive correlations being inordinately common among studies with lower precision. Using data for the AWC measure (21 studies) only to avoid a power inflation due to repeated measures, we assessed the presence of a potential publication bias with a funnel plot (see Figure 2) and the Egger's test for funnel plot asymmetry. The test revealed significant asymmetry for all the data (k = 31, z = 2.601, p = .009) with similar but non-significant trends for the 'published' effect sizes only (k = 13,

[1]It is also worth noting that we considered the different ways in which the SES construct was measured in individual studies (i.e. SES type: income-based, educational based, and composite) in order to be able to explore its value as moderator of the strength of the association between SES and experiences. The moderator was not significant, indicating that the correlation does not differ significantly as a function of the type of SES measure (see Supplementary Materials, Piot *et al.*, 2020).

https://doi.org/10.1017/S0305000921000441 Published online by Cambridge University Press

z = 1.757, p = .079) and the 'unpublished' effect sizes (k = 18, z = 1.353, p = .176). We used the trim-and-fill method to estimate the number of studies needed to symmetrize the plot based on all data, and to impute the effect size had these studies been available. This resulted in the addition of 12 imputed studies, leading to a weighted median effect size $r_z = .116$, SE = .031, CI = [.056; .177].

Discussion

Our systematic review of LENA™ measures of children's language experiences as a function of parental SES revealed that English-speaking countries provide the near totality of the data, with half of the effect sizes coming from the USA in particular. Thus, conclusions here may not generalize to other countries. Other than that, samples seemed quite diverse: the socioeconomic range was large, and on average half of the data came from Black, Asian, and Minority Ethnic populations – although it should be noted that ethnic characteristics were reported in less than half of the articles.

Our quantitative integration suggested the association between SES and LENA™ measures was small: in our model controlling for the type of LENA™ metric, age and publication type, the intercept was equal to z-transformed $r = .186$. Since funnel plots and Egger's tests were consistent with a publication bias in the AWC measures, we applied a correction for asymmetry, which resulted in an estimate of $r = .116$, which was still significant. Overall, these estimates align better with estimations of SES effects based on Sperry *et al.* (2019)'s data on overall speech quantities ($r = .182$) than either of their child-directed counts, which was $r = -.185$ for all adults, and $r = .609$ for the primary caregiver, which coincides with the fact that LENA™ AWC is based on all speech, and not just directed speech.

In fact, we had predicted the correlation strength to be strongest for CTC, which was not the case, with no significant difference in the correlation strength between this measure supposedly tagging interactive talk (CTC $r_z = .183$), and AWC ($r_z = .186$). Interestingly, our moderator analysis suggested that the estimate for CVC ($r_z = .094$) was significantly lower than that of the baseline AWC. These results suggest that the strongest associations between SES and language experience pertain to the overall near and clear adult speech, as well as to the back-and-forth interactions, with lower association strength for actual child speech.

In passing, we notice that the estimate for CVC is much lower than previous reports of SES association with word comprehension scores ($r_z = .26$, Scaff (2019)), suggesting that vocal production is less affected by SES differences than lexical development. This has sometimes been discussed for very early production measures, with significant effects of SES found for volubility, but not canonical proportion (Oller, Eilers, Basinger, Steffens & Urbano, 1995).

Finally we note several limitations. First, LENA™ measures capture meaningful variation in children's language experiences (Wang *et al.*, 2020) but touch upon conceptually distinct constructs, ranging from the child's own production (CVC) to overall speech in the child's input (AWC, which does not distinguish between child-directed and overheard talk). Although CTC has been viewed as an ideal measure of back-and-forth conversation (Romeo *et al.*, 2018), there are some open questions about its reliability (Cristia *et al.*, 2020), and its external validity, given its high correlation with CVC (for there to be a turn, there must be a child vocalization, leading to correlations between CVC and CTC above $r = .9$). It would be ideal for

https://doi.org/10.1017/S0305000921000441 Published online by Cambridge University Press

future generations of LENA or similar software to estimate child-directed speech separately from speech addressed to others, since this is likely to be an important distinction. Indeed, a preprint for a meta-analysis accepted as a registered report compares low- against mid- or high-SES groups in terms of child-directed word counts and overall word-counts, based mainly on studies relying on human annotations of short audio- or video-recordings (Dailey & Bergelson, 2021). They find a large effect estimated at Hedges' g = 0.69 ($r_z \sim .34$) for child-directed estimates, with a lower effect of about Hedges' g = 0.17 ($r_z \sim .09$) for overall estimates. Although those results may not translate to daylong recordings that are automatically analyzed, they do suggest that SES may affect child-directed speech quantities to a greater extent than overall quantities.

Returning to the limitations of our study, a second limitation is that this meta-analysis is only as good as the literature it is based on. Our data is drawn mainly from English-speaking countries, but the association between SES and language use in the home may vary across cultures and countries. Moreover, we found some evidence consistent with publication and reporting biases.

Finally, this study only measured the strength of the association, but cannot establish causality – which is extremely hard to assess, given that SES is a complex composite concept. Nevertheless, interventions that help extend the length of schooling and that are administered in a randomized control trial (e.g., Dupas, Duflo & Kremer, 2016) could be well placed to establish such causal chains by comparing mothers whose education was extended against control mothers.

In sum, it is important to accurately measure the potential strength of association between SES and children's early experiences, both for our theories of language acquisition and for potential implications in terms of social programs. We contribute to this goal by reviewing and integrating ecological data available in the current literature. Our results can be interpreted as being aligned with researchers arguing that we should look at language environments as a function of socioeconomic status, as there could be quantitative differences in the experiences afforded to young children (Golinkoff *et al.*, 2019), since we find significant correlations with SES even in these highly naturalistic data. However, the strength of the correlation is quite small.

Acknowledgments. AC acknowledges Agence Nationale de la Recherche (ANR-17-CE28-0007 LangAge, ANR-16-DATA-0004 ACLEW, ANR-14-CE30-0003 MechELex, ANR-17-EURE-0017); and the J. S. McDonnell Foundation Understanding Human Cognition Scholar Award.

The authors made the following contributions. Leonardo Piot: Conceptualization, Data entry, Analyses, Writing - Original Draft Preparation, Writing - Review & Editing; Naomi Havron: Writing - Review & Editing; Alejandrina Cristia: Conceptualization, Analyses, Writing - Review & Editing.

Our data is fully available in the corresponding OSF repository: https://osf.io/rw6ve/?view_only=eddcd33627d940c8ab51f5b7ef8e252d.

References

Adams, K. A., Marchman, V. A., Loi, E. C., Ashland, M. D., Fernald, A., & Feldman, H. M. (2018). Caregiver talk and medical risk as predictors of language outcomes in full term and preterm toddlers. *Child Development*, *89*(5), 1674–1690.

Beecher, C. C., & Van Pay, C. K. (2019). Small talk: A community research collaboration to increase parental provision of language to children. *Child & youth care forum*, 1–26. Springer.

Bergelson, E. (2017). *Bergelson seedlings home bank corpus.*

Brushe, M. E., Lynch, J. W., Reilly, S., Melhuish, E., & Brinkman, S. A. (2020). How many words are australian children hearing in the first year of life? *BMC Pediatrics*, *20*(1), 1–9.

Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, *10*(2), 159–199.

Christakis, D. A., Lowry, S. J., Goldberg, G., Violette, H., & Garrison, M. M. (2019). Assessment of a parent-child interaction intervention for language development in children. *JAMA Network Open*, *2* (6), e195738–e195738.

Cristia, A., Bulgarelli, F., & Bergelson, E. (2020). Accuracy of the language environment analysis system segmentation and metrics: A systematic review. *Journal of Speech, Language, and Hearing Research*, *63* (4), 1093–1105.

Dailey, S., & Bergelson, E. (2021). Language input to infants of different socioeconomic statuses: A quantitative meta-analysis. *Developmental Science, Registered report*.

d'Apice, K., Latham, R. M., & Stumm, S. von. (2019). A naturalistic home observational approach to children's language, cognition, and behavior. *Developmental Psychology*, *55*(7), 1414.

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

Dupas, P., Duflo, E., & Kremer, M. (2016). *Estimating the impact and cost-effectiveness of expanding secondary education in ghana*.

Dwyer, A., Jones, C., Davis, C., Kitamura, C., & Ching, T. Y. (2019). Maternal education influences australian infants' language experience from six months. *Infancy*, *24*(1), 90–100.

Ferjan-Ramírez, N., Lytle, S. R., & Kuhl, P. K. (2020). Parent coaching increases conversational turns and advances infant language development. *Proceedings of the National Academy of Sciences*, *117*(7), 3484–3491.

Flood, M. M. (2015). *Quality versus quantity: An investigation of the impact of home language and maternal education level on young children's vocabulary size*. UNIVERSITY OF WISCONSIN-MADISON.

Ganek, H., Smyth, R., Nixon, S., & Eriks-Brophy, A. (2018). Using the language environment analysis (lena) system to investigate cultural differences in conversational turn count. *Journal of Speech, Language, and Hearing Research*, *61*(9), 2246–2258.

Gilkerson, J., & Richards, J. A. (2008). *The lena natural language study*. Boulder, CO: LENA Foundation.

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.

Glick, P. C. (1976). Living arrangements of children and young adults. *Journal of Comparative Family Studies*, *7*(2), 321–333.

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

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young american children*. Paul H Brookes Publishing.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*(5), 1368–1378.

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.

Law, J., Charlton, J., & Rush, R. (2018). *Home-start early speech and language study: Phase 1 evaluation report*. Home-Start.

Lease-Johnson, E. (2018). *Redefining the word gap from a cumulative risk perspective* (PhD thesis). UNIVERSITY OF MINNESOTA.

Leung, C. Y., Hernandez, M. W., & Suskind, D. L. (2020). Enriching home language environment among families from low-ses backgrounds: A randomized controlled trial of a home visiting curriculum. *Early Childhood Research Quarterly*, *50*, 24–35.

McGillion, M., Pine, J. M., Herbert, J. S., & Matthews, D. (2017). A randomised controlled trial to test the effect of promoting caregiver contingent talk on language development in infants from diverse socioeconomic status backgrounds. *Journal of Child Psychology and Psychiatry*, *58*(10), 1122–1131.

Merz, E. C., Maskus, E. A., Melvin, S. A., He, X., & Noble, K. G. (2020). Socioeconomic disparities in language input are associated with children's language-related brain structure and reading skills. *Child Development*, *91*(3), 846–860.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *PLOS Medicine*, *6*(7), 1–6. https://doi.org/10.1371/journal.pmed.1000097

Orena, A. J., Byers-Heinlein, K., & Polka, L. (2019). Reliability of the language environment analysis recording system in analyzing french–english bilingual speech. *Journal of Speech, Language, and Hearing Research*, *62*(7), 2491–2500.

Oller, D. K., Eilers, R., Basinger, D., Steffens, M., & Urbano, R. (1995). Poverty and speech precursor development. *First Language*, *15*(2), 167–187.

Piot, L., Havron, N., & Cristia, A. (2020). *Socioeconomic status correlates with measures of Language ENvironment Analysis (LENA) system: a meta-analysis*. Retrieved from https://osf.io/rw6ve/?view_only=eddcd33627d940c8ab51f5b7ef8e252d

R Core Team (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Romeo, R. R., Leonard, J. A., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2019). *Neural correlates of the "word gap": Children's language experience is associated with brain structure and function*. SRCD.

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

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). Publication bias in meta-analysis. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, 1–7.

Scaff, C. (2019). *Beyond WEIRD: An interdisciplinary approach to language acquisition* (PhD thesis).

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019). Reexamining the verbal environments of children from different socioeconomic backgrounds. *Child Development*, *90*(4), 1303–1318.

Sultana, N., Wong, L. L., & Purdy, S. C. (2020). Natural language input: Maternal education, socioeconomic deprivation, and language outcomes in typically developing children. *Language, Speech, and Hearing Services in Schools*, *51*(4), 1049–1070.

Swanson, M. R., Donovan, K., Paterson, S., Wolff, J. J., Parish-Morris, J., Meera, S. S., Watson, L. R., Estes, A. M., Marrus, N., Elison, J. T., Shen, M. D., McNeilly, H. B., MacIntyre, L., Zwaigenbaum, L., St John, T., Botteron, K., Dager, S., Piven, J., & IBIS Network (2019). Early language exposure supports later language skills in infants with and without autism. *Autism Research*, *12*(12), 1784–1795. https://doi.org/10.1002/aur.2163

VanDam, M. (2018). *Cougar homebank corpus*.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. Retrieved from https://www.jstatsoft.org/v36/i03/

Wang, Y., Williams, R., Dilley, L., & Houston, D. M. (2020). A meta-analysis of the predictability of lena™ automated measures for child language development. *Developmental Review*, *57*, 100921.

Warlaumont, A. S., Pretzer, G. M., Walle, E., Mendoza, S., & Lopez, L. (2016). *Warlaumont homebank corpus*.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152.

Zegiob, L. E., Arnold, S., & Forehand, R. (1975). An examination of observer effects in parent-child interactions. *Child Development*, *46*(2), 509–512. Retrieved from http://www.jstor.org/stable/1128149

Cite this article: Piot L, Havron N, Cristia A (2022). Socioeconomic status correlates with measures of Language Environment Analysis (LENA) system: a meta-analysis. *Journal of Child Language* **49**, 1037–1051. https://doi.org/10.1017/S0305000921000441