

# The use of Language ENvironment Analysis in autism research: A systematic review

Orla C Putnam<sup>1\*</sup> , Jennifer E Markfeld<sup>2\*</sup> ,  
Sarah Towner Wright<sup>1</sup> , Jacob I Feldman<sup>2,3</sup> ,  
Jessica Goldblum<sup>4</sup> , Maia Karpinsky<sup>1</sup>, Amanda J Neal<sup>1</sup>,  
Meghan R Swanson<sup>5</sup> and Clare Harrop<sup>1</sup> 

Autism

1–23

© The Author(s) 2024



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/13623613241290072

[journals.sagepub.com/home/aut](https://journals.sagepub.com/home/aut)

## Abstract

The Language ENvironment Analysis (LENA) system has been used increasingly in research to record and evaluate the everyday speech of autistic children and their families. However, it is unclear how researchers are using LENA and whether the system is well-suited for work with autistic individuals. The purpose of this systematic review is to summarize the use of LENA in autism research, to highlight the strengths and limitations of the system as identified by researchers, and to provide recommendations for future research and clinical use. Forty-two studies that used LENA with samples of autistic children were identified through a systematic database search. Researchers using LENA in autism did so across a variety of ages, settings, and analytical approaches. Most studies used LENA within recommended guidelines. The most common purpose of using LENA was for exploratory research. Noted strengths of the LENA system included ecological validity, cost-effectiveness, and timely clinical feedback. Limitations included lower rates of speaker identification compared to human coders and limited information regarding speech context and language development. This systematic review provides key insights into the methods surrounding LENA use in autism research and serves to inform researchers and clinicians on best practices for future use with this technology.

## Lay abstract

In research, language ability has historically been measured using structured tasks in laboratory settings. In recent years, there has been a growing emphasis on the need to instead capture language ability in an individual's natural setting (i.e. through social interaction or in their home). Considering natural language may be particularly important for the autistic population, as an autistic child's language ability can be very different depending on the setting. One common tool for capturing natural language is the LENA recording system, which takes audio recordings over long periods of time and provides estimates of children's and caregivers' speech. The purpose of this systematic review is to summarize the use of LENA in autism research, to highlight the strengths and limitations of the system as identified by researchers, and to provide recommendations for future research and clinical use. We identified 42 autism studies that used LENA in a variety of ways and settings. Most studies used LENA within the guidelines put forth by its creators, and it was most commonly used to understand speech or speech development for autistic children. LENA is a useful tool for clinicians and caregivers to gain some insights into child speech, but those considering using it should be aware of concerns about its accuracy and limitations about the information it provides. In this review, we supplement the official LENA guidelines with specific suggestions for use with the autistic population.

## Keywords

autism, communication and language, technology

<sup>1</sup>The University of North Carolina at Chapel Hill, USA

<sup>2</sup>Vanderbilt University, USA

<sup>3</sup>Vanderbilt University Medical Center, USA

<sup>4</sup>3-C Institute for Social Development, USA

<sup>5</sup>University of Minnesota, USA

\*Orla C. Putnam and Jennifer E. Markfeld contributed equally to this work.

## Corresponding author:

Orla C Putnam, Division of Speech and Hearing Sciences, The University of North Carolina at Chapel Hill, Bondurant Hall, 321 S Columbia St., Chapel Hill, NC 27514, USA.  
Email: [orla\\_putnam@med.unc.edu](mailto:orla_putnam@med.unc.edu).

Despite its removal from the official diagnostic requirements, language continues to be a primary area of challenge for many autistic individuals (Georgiou & Spanoudis, 2021; Tager-Flusberg, 2006; Tager-Flusberg et al., 2009; Wittke et al., 2017) and remains an important area of investigation. Differences in language development are evident at an early age in autistic children, increasing over time relative to neurotypical children (Messinger et al., 2013). Approximately 30% of autistic individuals are minimally speaking (Tager-Flusberg & Kasari, 2013), and concerns about language development are the leading cause for referrals for an autism evaluation (Dillon et al., 2021; Harrop et al., 2021; Mattheis et al., 2017; McCormick et al., 2020). Many autistic individuals receive speech-language services for language and communication abilities (American Speech-Language-Hearing Association, n.d.), as these abilities are highly predictive of later social, educational, and vocational success for people on the autism spectrum (e.g. Billstedt et al., 2005; Eaves & Ho, 2008). In addition, language interventions and speech services have been promoted as a high-need area of therapy by autistic adults (Benevides et al., 2020).

Given that language development is so heterogeneous in autism, the ways that language is measured can inform treatment recommendations and how language abilities are characterized in this population. Research on language and social communication in autism has historically been conducted in controlled laboratory settings using structured language assessments or parent-report measures. However, observations of natural, unstructured language have been prioritized by researchers over the last two decades to capture elements of language that may not be present in structured assessment settings (Barokova & Tager-Flusberg, 2020). These short, observational periods of unstructured language production and use typically take place over a 5-to-10-min recorded session of free-play or conversation with a caregiver or researcher (Channell et al., 2018; La Valle et al., 2024). From these samples of language use, researchers and clinicians can derive indices of language such as the rate of speech and conversations, as well as grammatical complexity and grammar and/or speech errors (e.g. Channell et al., 2018; La Valle et al., 2024; Spencer et al., 2023; Winters et al., 2022).

While this extant body of research provides a rich foundation on which the field has characterized language in autism, there has been a growing emphasis on the idea that a child's communication over a short period in an unfamiliar environment may not be fully representative of their natural language use (Bergelson et al., 2019; Woynaroski et al., 2017). This emphasis has driven researchers to develop innovative ways to capture speech and language in naturalistic, ecologically valid settings. This review focuses on one research tool, the Language ENvironment Analysis (LENA) system, and its use in autism research.

## The use of LENA to study language in autism

The push toward the measurement of a child's natural language environment has resulted in the increased use of *daylong recording technology*: Non-invasive recording devices that can record audio for considerable lengths of time in everyday environments without the presence of an observer (Cychosz & Cristia, 2021). One such device, LENA, is widely used by researchers and clinicians (Gilkerson et al., 2017). The LENA device allows for up to 16 h of audio recordings, and accompanying software algorithmically categorizes types of vocalizations and other sounds that occur (e.g. child vocalizations, adult vocalizations, nonspeech sounds). The LENA software produces raw counts and summary statistics of each variable of interest (Gilkerson et al., 2017). These automated calculations allow researchers and clinicians to gain information about a child's language environment without the time and cost of manual transcription and annotation. Although LENA can provide estimates of child speech and the home language environment, it is important to note that it cannot capture more fine-grained linguistic information such as linguistic diversity. In addition, the algorithm normed by the LENA Foundation was trained on a small sample of 94 monolingual, neurotypical, North American children (Gilkerson & Richards, 2020).

Since its development, and despite LENA's processing algorithms being trained on relatively small sample of neurotypical children, LENA has been advertised as a method of early identification of autism (Gilkerson et al., 2017). Researchers at the LENA Foundation have created the Automatic Autism Screen (AAS), with the goal of objectively distinguishing autistic children from neurotypical children and children with language delay using acoustic information derived from daylong recordings (Richards et al., 2010). However, the AAS was normed on 232 autistic children from ages 16–48 months, which brings into question its ability to predict autism diagnosis earlier than clinicians (Oller et al., 2010). The AAS study reported sensitivity and specificity values for differentiating between neurotypical, autistic, and language-delayed children based on vocal features. When differentiating between autistic and neurotypical the AAS classifier yielded sensitivity and specificity values between .82 and .84, which are values generally considered acceptable for diagnostic purposes; however, the classifier was unable to clearly differentiating between autistic and language-delayed or typical and language-delayed. (Oller et al., 2010). In addition, the AAS has not been independently validated, nor is it currently publicly available, and therefore may have limited utility for researchers and clinicians.

## **LENA foundation guidelines for use of LENA**

Given that LENA is widely used in research, the LENA Foundation has established guidelines to aid in the proper use of the device and accompanying software. First, they recommend that LENA be used with children 2–48 months of age, and within the home environment, as these are the ages and settings on which the device's algorithm was trained (Gilkerson & Richards, 2020). In addition, they suggest that 16 recording hours, LENA's maximum recording length, be collected to obtain stable estimates of talk throughout the day. However, it should be noted that a 16-h recording is likely to capture periods when young children are asleep (Paruthi et al., 2016). The LENA Foundation suggests that if fewer than 16 h of data are collected, users should randomly select continuous 10-min intervals to analyze and should sample “varying levels of speech activity” if recording for less than an hour (Gilkerson & Richards, 2020). In addition, recordings should not be “restricted to selections from short recordings made in controlled environments” (Gilkerson & Richards, 2020).

## **Clarity around use of LENA in autism research**

Despite the widespread use of LENA in autism research over the past decade, it is unclear whether researchers are following the guidelines put forth by the LENA Foundation or whether they are using LENA in unique ways. As researchers and clinicians may seek to replicate and/or apply the methods implemented by published studies in the field, it is important to systematically evaluate such methods with respect to the LENA guidelines. In addition, given that LENA's algorithm was not normed on autistic children, there is still much to be learned about the potential validity and reliability concerns of using LENA in a population where speech and language development is more heterogeneous. Understanding how LENA is being used in autism research is critical, particularly given the need to establish ecologically valid measures of speech production in autism and to capture naturalistic language in a cost-effective manner (e.g. Cychosz & Cristia, 2021; McDaniel, Yoder, Estes, & Rogers, 2020). Researchers and clinicians may be inclined to use LENA within their work due to its unobtrusive properties, automated software, and precedence in the literature for using LENA, however the system may not be the best fit for all intended uses.

### **Purpose**

Existing reviews have examined aspects of language in autism such as parent verbal responsiveness (Edmunds

et al., 2019), links to neuroimaging (Butler et al., 2020), and overall intervention efficacy (Sandbank et al., 2020). Past reviews of LENA use in neurotypical children have focused on ethics (Cychosz et al., 2020), validity of automated metrics (Cristia et al., 2020; Wang et al., 2017, 2020) and the overall use of the device in research (Greenwood et al., 2018). However, there is currently no synthesized information, to our knowledge, on how researchers are using LENA specifically to characterize autistic children's speech input and output.

The purpose of this systematic review, therefore, is to (a) provide researchers and clinicians with a summary of how LENA has been used in autism research, (b) summarize the strengths and limitations of using LENA with this population as identified by researchers, and (c) provide recommendations and considerations for LENA use with the autistic population.

## **Method**

This review was registered online with PROSPERO (registration number: CRD42021255947). We followed the PRIMSA checklist in writing and reporting this systematic review. A trained health sciences librarian STW performed our comprehensive electronic search of publications using the following databases: PubMed, Cumulative Index to Nursing and Allied Health Literature via EBSCO, EMBASE.com, Scopus, PsycINFO via EBSCO, and ERIC via EBSCO. These databases were initially searched in June 2021 and were searched again in March 2023. Our search was not restricted by language.

All database results were collected from January 2008 (after the creation of the LENA device) to March 2023. Search terms were used to retrieve articles addressing the main concepts of the search strategy: LENA/language counts and autism. The search strategy was conducted using a combination of text word searching and the use of subject headings/thesaurus terms, if applicable. The exact search strategy used in each of the electronic databases is reported in Supplemental Information S1. Results were downloaded to EndNote and duplicates removed. All references were uploaded to Covidence Systematic Review software (<https://www.covidence.org>), a web-based tool designed to facilitate and track each step of the abstraction and review process.

### **Study selection**

Abstracts identified by the search strategy were screened twice and independently by O.C.P. and J.G., who excluded records that did not mention language, autism, or similar broad search criteria. Full texts of the remaining records were independently screened by O.C.P., J.E.M., and J.G., with two people screening each full text. At both stages of the screening process, CH resolved disagreements when

needed. All full text articles included in the review were additionally screened using forward and backward citation searching to find any additional relevant articles. Article inclusion criteria included: contain primary data (i.e. not a review article), use the LENA device, and include a group with a diagnosis of autism or a “high likelihood” group. We defined a high likelihood group as any sample that was at an increased likelihood to go on to have an autism diagnosis, such as infant siblings and young children identified to be high likelihood via community screenings (e.g. Messinger et al., 2013; Watson & Crais, 2013). Because our coding scheme was aimed at understanding how LENA has been used in autism research, including whether or not researchers are using LENA within the age groups recommended by the LENA Foundation guidelines, we did not specify a participant age group in our inclusion criteria. To avoid publication bias (Paez, 2017), gray literature (e.g. conference papers, honors, masters, doctoral theses) that were found in our search results were not excluded in our full-text review.

We conducted a separate gray-literature search of proceedings and abstract books of major conferences relevant to language, autism, or both (American Speech and Hearing Association, International Society for Autism Research, Society for Research in Child Development, Symposium for Research in Child Language Disorders) via hand-searching in February 2022 and January 2023 using our key search terms and dating back to 2008 when available. The resulting abstracts were double screened independently by O.C.P. and J.E.M.. All authors of abstracts deemed eligible were contacted for full texts and these abstracts or associated materials were further assessed for eligibility by O.C.P., J.E.M., and J.G. After the screening process, 42 studies met eligibility criteria for inclusion in the current review (31 peer-reviewed journal articles; three doctoral dissertations; eight conference papers, posters, and presentations). See Figure 1 for full screening information.

## Coding

A 34-item deductive coding scheme (see Supplementary Information S2 for list of items) was developed by the first two authors, with MRS and CH providing feedback on early versions prior to coding. Data on participant demographics, study purpose, information pertaining to the use of LENA (e.g. variables used from LENA software, average recording length, how the data were used), and identified strengths and limitations of using LENA were extracted from each study. An inductive coding method and semantic approach was used to code and analyze the strengths and limitations noted: Any explicitly-stated strengths, limitations, and recommendations for LENA use were flagged during the initial study coding and agreed upon by the coders, and later grouped into themes that

developed as the data were analyzed (Braun & Clarke, 2006; Saldaña, 2016). Authors O.C.P., J.E.M., M.K., and A.N. coded each article independently, with the exception of Markfeld, Feldman, Bordman, et al. (2023), which was independently coded by O.C.P. and J.G. Any disagreement between the two sets of codes were resolved via discrepancy discussions.

## Quality assessment

Study quality was assessed for full-text articles (e.g. not conference posters) using the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018). See full information in Supplemental Information S3.

## Community involvement

Autism community members were not involved with this systematic review.

## Results

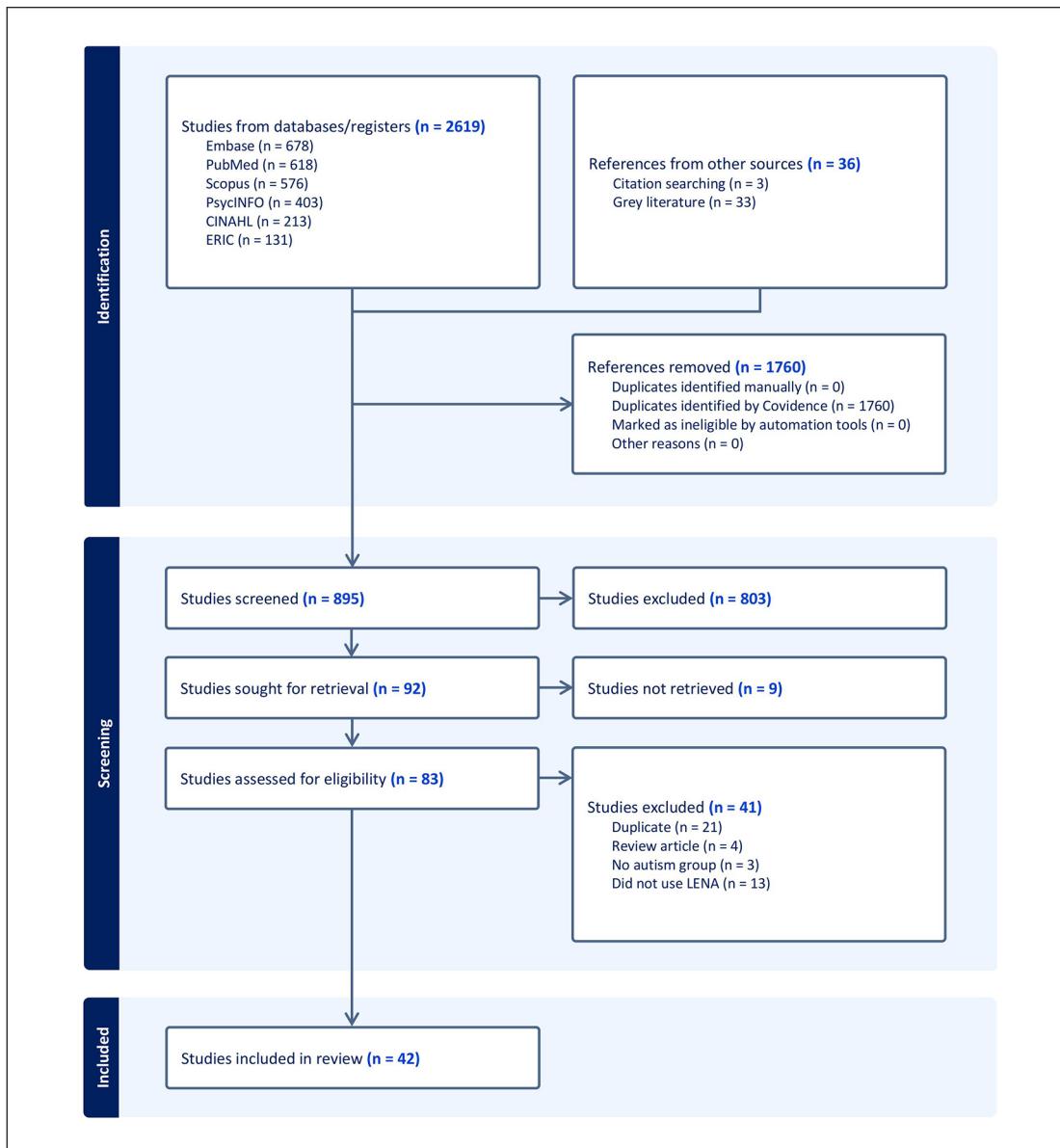
The aim of the current review is to summarize how LENA has been used in autism research and what strengths and limitations of the system have been identified by researchers. The purpose, participant characteristics, location and recording settings, and analytical approach of the forty-two included studies are detailed below, followed by the qualitatively coded strengths and limitations of LENA as described by researchers. See Table 1 for details on each included study.

## Quality assessment of included studies

Coding via the MMAT (Hong et al., 2018) suggests that full-text studies were overall high quality with some exceptions. Sixteen studies did not report the race and/or sex of participants; therefore, it is not possible to evaluate whether these studies are representative of all autistic individuals (e.g. Steinbrenner et al., 2022). See Supplemental Information S3 for detailed information.

## Research purposes

The majority of the studies ( $k=26$ ) had *exploratory* purposes, using LENA to observe and describe speech. Eight of these studies fell under both *exploratory* and another purpose category. Within exploratory studies, the purpose of LENA use was to understand language development and outcomes in autistic children (Bak et al., 2019; Fleurissant, 2017; Ghai, 2013; Markfeld, Feldman, Bordman, et al., 2023; McDaniel, Yoder, Estes, & Rogers, 2020; Seidl et al., 2018; Swanson et al., 2019; Woynaroski et al., 2017; Xu et al., 2014) and to better understand the language environment of autistic children (Burgess et al., 2013; Dykstra et al., 2013; Ferguson et al., 2020; Fragoso,



**Figure 1.** Prisma flow diagram showing study selection.

2016; Irvin et al., 2013; Little et al., 2019; Warren et al., 2010). LENA was also used to compare vocalizations of autistic children and children with a high likelihood of autism to non-autistic/low-likelihood control groups (Fasano et al., 2021; Ghai, 2013; Meera, 2019; Seidl et al., 2018; Swanson et al., 2019; Warlaumont et al., 2014; Warren et al., 2010). Finally, LENA was also used to examine the associations between LENA metrics of speech and autistic traits (Dykstra et al., 2013; Ferguson et al., 2020; Irvin et al., 2013; Moffitt et al., 2022; Rankine et al., 2017; Sulek et al., 2022) and clinical measures of language (Dykstra et al., 2013; McDaniel, 2017; Sabatos-DeVito, 2016; Yoder et al., 2013).

Studies using LENA for *measure creation and/or validation* ( $k=13$ ; 6 in combination with exploratory) either

tested the efficacy of LENA when used with age ranges or amount of recording time outside the existing guidelines of use (Bredin-Oja et al., 2018; Jones et al., 2019; Yoder et al., 2013), tested the validity of LENA metrics with an autism-specific sample (Bruyneel et al., 2021; McDaniel, Yoder, Estes, & Rogers, 2020; Oller, 2015; Oller et al., 2010; Pawar et al., 2017; Rankine et al., 2017; Sabatos-DeVito, 2016; Woynaroski et al., 2017; Xu et al., 2014), or tested the validity of a new speech metric using LENA equipment (Harbison et al., 2018; McDaniel, 2017; Yoder et al., 2013).

Other studies used LENA to test the efficacy of an *intervention* ( $k=9$ ; 1 in combination with exploratory), or for the *early detection/screening* ( $k=2$ , 1 in combination with exploratory) of autism in children identified to be at high likelihood of an autism diagnosis.

**Table 1.** Study information.

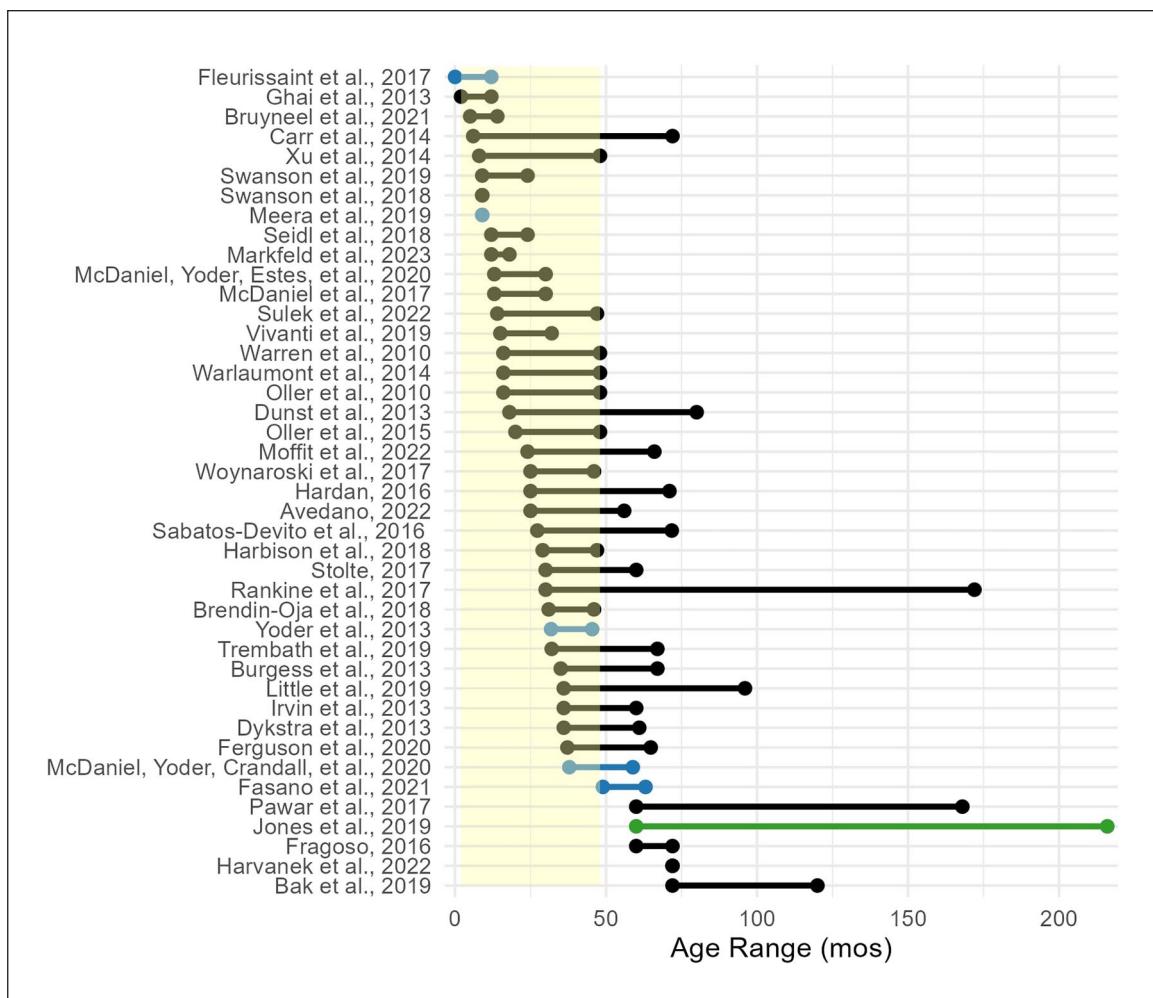
Study	Purpose of study	Sample size	Average age	Age range	Percent male	Home language	Diagnostic status
Avendano (2022)	Intervention	31	41.55 mo	25–56 mo	77	English	Autism
Bak et al. (2019)	Exploratory	11	7.35 years	6–10 years	100	NR	Autism
Bredin-Oja et al. (2018)	Measure Creation/ Validation	6	36 mo	31–46 mo	100	English	Autism, IDD, & DS groups
Bruyneel et al. (2021)	Measure creation/ validation	12	5.59, 10.47, & 14.30 mo	5, 10, & 14 mo	42	Dutch	Infant Siblings
Burgess et al. (2013)	Exploratory	10	54.7 mo	35–67 mo	100	English	Autism
Carr et al. (2014)	Early detection/ Screening	83	27.6 mo	6–72 mo	NR	NR	Community Risk
Dunst et al. (2013)	Intervention	11	45.36 mo	18–80 mo	82	NR	Autism
Dykstra et al. (2013)	Exploratory	40	48.3 mo	36–61 mo	80	NR	Autism
Fasano et al. (2021)	Exploratory	56	50.14 mo	NR	61	English	Autism
Ferguson et al. (2020)	Exploratory	53	4.1 years	3.1–5.4 years	78	NR	Autism
Fleurissement et al. (2017)	Exploratory	74	NR	NR	NR	NR	Infant Siblings; NA group
Fragoso (2016)	Exploratory	6	NR	5–6 years	NR	English & Spanish	Autism
Ghai et al. (2013)	Exploratory	8	NR	2–12 mo	NR	NR	Infant Siblings; NA group
Harbison et al. (2018)	Measure creation/ validation	21	41 mo	29–47 mo	86	NR	Autism
Hardan (2016)	Intervention	37	46.3 mo	25–71 mo	86	NR	Autism
Harvánek et al. (2022)	Intervention	1	6 years	6 years	0	NR	Autism
Irvin et al. (2013)	Exploratory	67	NR	3–5 years	79	NR	Autism
Jones et al. (2019)	Measure creation / validation	24	NR	5–18 years	79	NR	Autism
Little et al. (2019)	Exploratory	10	73.6 mo autism; 63 mo NA	3–8 years	NR	English	Autism group; NA group
Markfeld, Feldman, Bordman, et al. (2023)	Exploratory	50	13.71 mo HL; 14.05 mo LL	12–18 mo & 21–27 months	53.6 HL; 50 LL	English	Infant Siblings
McDaniel (2017)	Exploratory, Measure Creation/Validation	68	NR	13–30 mo	NR	NR	Autism
McDaniel, Yoder, Crandall, et al. (2020)	Exploratory, Measure Creation/Validation	40	49.8 mo	NR	88	English	Autism
McDaniel, Yoder, Estes, & Rogers (2020)	Measure creation/ validation, Exploratory	87	23.42 mo	13–30 mo	76	English	Autism
Meera (2019)	Exploratory	72	9 mo	NR	NR	NR	Infant Siblings; NA group
Moffitt et al. (2022)	Exploratory	66	39.97 mo	24–66 mo	74	NR	Autism group; Developmental Delay group
Oller et al. (2010)	Measure creation/ validation	232	35.93 mo autism; 28.18 mo NA;	NR	45 NT; 67 Lang Delay;	NR	Autism
Oller et al. (2015)	Measure Creation/ Validation	232	27.96 Lang Delay	NR	83 autism NR	NR	Autism
Pawar et al. (2017)	Measure creation/ validation	20	NR	5–14 years	NR	NR	Autism

(Continued)

**Table 1.** (Continued)

Study	Purpose of study	Sample size	Average age	Age range	Percent male	Home language	Diagnostic status
Rankine et al. (2017)	Exploratory, Measure creation/validation	18	76.92	30–172 mo	44	NR	Autism
Sabatos-DeVito (2016)	Exploratory	25	4.47 years	2.27–5.98 years	NR	NR	Autism
Seidl et al. (2018)	Exploratory	36	NR	12–24 mo	64	English	Infant Siblings
Stolte (2017)	Intervention	3	NR	30–60 mo	66%	English, Punjabi, & Urdu	Autism
Sulek et al. (2022)	Intervention, Exploratory	99	31.55 mo	14–47 mo	70.70	NR	Autism
Swanson et al. (2018)	Exploratory	59	9.45 mo HL; 9.53 mo LL	9 mo	65 HL; 47 LL	English	Infant Siblings
Swanson et al. (2019)	Exploratory	96	25.34 mo HL+; 24.71 mo HL-; 25.14 mo LL-	9–24 mo	71.43 HL+; 58.70 HL-; 58.33 LL-	NR	Autism
Trembath et al. (2019)	Intervention	23	49.6 mo	32–67 mo	74	English & Other	Autism
Vivanti et al. (2019)	Intervention	44	24.73 & 26.36 mo	15–32 mo	61	NR	Autism
Warlumont et al. (2014)	Exploratory	183	NR	16–48 mo	83 autism; 45 NA	NR	Autism group; NA group
Warren et al. (2010)	Exploratory	104	mo NA	autism; 8–48	mo NA	English	Autism
Wojnaroski et al. (2017)	Exploratory, Measure creation/validation	20	32.9 & 30 mo	16–48 mo	85	English	Autism
Xu et al. (2014)	Exploratory, Measure creation/validation	226	37.9 mo	25–46 mo	85%	English	Autism
Yoder et al. (2013)	Exploratory, Measure creation/validation	69	NR	8–48 mo	NR	NR	Autism group; NA group
			38.6 mo & 28.1 mo	NR	82 autism; 48 NA	NR	Autism group; NA group

Note. Italicized rows indicate studies from the gray literature search. NR = not reported; DS = Down Syndrome; IDD = Intellectual & Developmental Disability; HL = High-likelihood; LL = Low-likelihood; HL+ = High-likelihood with an autism diagnosis; HL- = High-likelihood without an autism diagnosis; NA = non-autistic.



**Figure 2.** Depiction of age ranges used across studies. The yellow highlighted area of the figure shows the age range that the LENA Foundation recommends for use of the software (i.e. 2–48 months). The study in green (Jones et al., 2019) purposefully used LENA outside of the recommended age range. Blue are studies where age range was estimated, either via mean and standard deviation of child age (i.e. Fasano et al., 2021; McDaniel, Yoder, Crandall, et al., 2020; Meera, 2019; Yoder et al., 2013), or via descriptive text (i.e. Fleurissant, 2017).

### Characteristics of participants in included studies

The sample sizes across studies ranged from 1 to 232 participants ( $M_{sample} = 57.92$ ,  $SD_{sample} = 60.32$ ).

**Age of participants.** Age ranges of participants ranged from infancy (2 months) to adolescence (Figure 2). Average ages for each study are reported in Table 1. Twenty-one studies included participants that were within the recommended LENA age guidelines (2 to 48 months), and 15 included some, but not all, participants in this range. Of the six studies that did not follow this guideline, one study (Jones et al., 2019) was conducted with the purpose of evaluating LENA outside the recommended age range.

**Participant demographics.** Of the studies that reported participant sex ( $k=31$ ), participants were overwhelmingly male ( $M_{sex} = 69.4\%$  male,  $SD_{sex} = 19.67\%$  male). One study (Harvánék et al., 2022) was a case study with one autistic female; the remaining studies reported that 44% to 100% of their samples were male. Of the studies that reported participant race ( $k=16$ ), participants were overwhelmingly white, ranging from 11% to 100% across studies ( $M_{race} = 66.65\%$  white,  $SD_{race} = 23.1\%$  white). Of the 16 studies that reported the languages spoken in recording environments, 12 reported the home language as English, three reported English and at least one other language (Bruyneel et al., 2021; Fragoso, 2016; Trembath et al., 2019). Bruyneel et al. (2021) featured Dutch speakers, and was the only study conducted with the purpose of validating LENA in another language.

**Diagnosis.** Most studies using LENA for autism research included participants who had an existing autism diagnosis ( $k=31$ ). Of these, participants in four studies were specified to be pre- or minimally verbal (Bak et al., 2019; Bredin-Oja et al., 2018; Harvánek et al., 2022; Jones et al., 2019). Eleven other studies included additional children with other diagnoses (e.g. language delay, Down syndrome), and 10 studies included a non-autistic (often termed “typically developing”) comparison group. Nine studies focused on or included siblings of autistic children. One study (Carr et al., 2014) focused on screening Deaf and hard of hearing children who were being evaluated for autism.

### **Location and recording settings**

Information pertaining to the specific use of LENA is detailed in Table 2.

**Location.** Three studies did not report where LENA recordings took place. Among those that did, LENA was used in home ( $k=23$ ), school ( $k=6$ ), and clinic ( $k=4$ ) settings. Six studies reported LENA use across multiple settings. The recording location was central to the aims of seven studies, which examined LENA use in school, intervention, and/or community settings (Burgess et al., 2013; Dykstra et al., 2013; Fasano et al., 2021; Ferguson et al., 2020; Jones et al., 2019; Little et al., 2019).

**Length & frequency of recording.** Nineteen of the studies used the recommended “all-day” recording length of 16-hours for their analyses: Three of these studies (Harbison et al., 2018; Markfeld, Feldman, Bordman, et al., 2023; Woynaroski et al., 2017) collected two all-day recordings for analysis, and another (Yoder et al., 2013), collected three. Eighteen studies used a structured and/or pre-determined recording time (e.g. the length of a school day or a set number of recording hours). This recording time ranged from 5 min to 12 h. Two studies reported analyzing 5 minute intervals of time rather than the entire recording either for transcription or coding of vocalizations (Bruyneel et al., 2021; Meera, 2019). The remaining 23 studies, representing over half, did not report the average length of their recordings or their sampling procedure (e.g. how they selected units for analysis). No studies discussed the potential influence of varying lengths of waking and sleeping hours in their sample. This may limit findings related to variables that do not inherently control for recording length.

### **Analysis of LENA data collected in included studies**

**LENA variables used.** Thirty studies utilized the built-in LENA software variables for analysis: Child Vocalization Count (an estimate of the number of times a child produces speech-like vocalizations) was the most widely

used ( $k=25$ ). Eighteen studies used Adult Word Count (an estimate of the number of words produced by adult males and females in the child’s environment) and Conversational Turn Count (an estimate of the number of alterations between the child and adult in the child’s environment that are fewer than five seconds apart). Other variables used included Male/Female Adult Near, which are estimates of the number of adult words produced near the target child, categorized into sexes based on fundamental frequency (Gilkerson & Richards, 2020); Automated Vocal Analysis, which is an estimate of child expressive language based on quantifying speech-like sounds in child vocalizations (Richards et al., 2008); and Nonspeech Sounds, which are instances of burps, cries, and other nonspeech sounds from the target child (Gilkerson & Richards, 2020).

Researchers in nine studies conducted analyses using variables not directly provided by the commercial LENA software (e.g. custom or extant software/algorithms). Common variables derived from extant methods included the Reciprocal Vocal Contingency (RVC) score, which is an index of caregiver-child vocal reciprocity (Harbison et al., 2018); the Average Count Per Utterance (ACPU), which estimates child productions of consonants, vowels, and nonspeech sounds within utterances that are labeled by LENA as being produced by the target child; and the Infraphonological Vocal Development (IVD) variable, which quantifies vocal complexity based on detecting speech-like acoustic parameters in child speech (Xu et al., 2014).

**Human transcription and coding.** Researchers in fourteen studies manually transcribed and/or coded audio recordings collected using LENA. Two studies utilized human transcription alongside output from the LENA system as part of their analyses. Burgess et al., (2013) transcribed the type of adult words directed at autistic children using audio segments that LENA identified as having the highest Adult Word Count values. Similarly, Fragoso (2016) transcribed adult vocalizations in bilingual households, coding them as English or Spanish to supplement LENA-generated vocalization counts. Three studies reported human transcription and coding as their only analysis, opting not to use LENA-derived indices of speech (Ferguson et al., 2020; Ghai, 2013; Meera, 2019). One such study (Ferguson et al., 2020) provided a rationale for this: Their pilot study revealed that child speech recorded in a school environment was more accurately identified by humans than by LENA, leading the team to use LENA solely for its recording capacity.

Eight studies compared their human transcriptions and codes to LENA output and reported on the findings. The purpose of five of these studies (Bredin-Oja et al., 2018; Bruyneel et al., 2021; Jones et al., 2019; Pawar et al., 2017; Rankine et al., 2017) was to evaluate the

**Table 2.** LENA sampling and analysis.

Study	LENA Information			Sampling Method	Length of Recording (average)	Human Transcription
	LENA Variables	Non-LENA Variables	Recording Setting			
Avendano (2022)	CTC, CVC	—	Clinic	Pre-determined time	6.4 h	N
Bak et al. (2019)	CTC, CVC	—	School	Pre-determined time	6.15 h a day	Y
Bredin-Oja et al. (2018)	CTC, CVC	—	NR	Pre-determined time	23 min	Y
Bruynseel et al. (2021)	CTC, AWC, CVC	—	Home	Pre-determined time	NR	Y
Burgess et al. (2013)	AWC	—	Home & School	Interval	Range 2.5 to 15.4 h	Y
Carr et al. (2014)	LLAS	—	NR	Peak speaking times	NR	Y
Dunst et al. (2013)	CVC	—	NR	Pre-determined time	NR—ranged from 5 to 27 min	N
Dykstra et al. (2013)	CTC, AWC, CTC	—	School	Pre-determined time	5 h 18 min	N
Fasano et al. (2021)	CVC	—	School	Peak speaking times	~2.5 h	Y
Ferguson et al. (2020)	—	—	School	Pre-determined time	NR	Y
Fleurissant et al. (2017)	CTC, AWC, CVC	—	Home	NR	N	N
Fragoso (2016)	AWC, CTC, CVC	—	Home	Pre-determined time	NR	Y
Ghai et al. (2013)	—	—	Home	NR	Y	N
Harbison et al. (2018)	RVC	—	Home	NR	N	N
Hardan (2016)	CVC, MAN/FAN	—	Home	Pre-determined time	1 h avg	N
Harvánek et al. (2022)	CTC, AWC, CVC	—	Home	Pre-determined time	~2–3 h	N
Irvin et al. (2013)	AWC	—	School	Pre-determined time	168 min	N
Jones et al. (2019)	CVC	—	Home, Clinic, Lab	Pre-determined time, Interval	NR	Y
Little et al. (2019)	CTC, AWC, CVC	—	Home, Community Settings	All day	NR	N
Markfeld, Feldman, Bordman et al. (2023)	AWC, MAN/FAN	—	Home	All day	NR	N
McDaniel (2017)	—	RVC	Home	Pre-determined time	NR	N
McDaniel, Yoder, Crandall, et al. (2020)	—	RVC	Home	All day	15 h, 25 min, 26 seconds for intervention group; 15 h, 2 min for control group	N
McDaniel, Yoder, Estes, & Rogers (2020)	AQV	APCU-C + V	Home	All day	NR	N
Meera (2019)	—	—	Home	Interval	NR	Y
Moffit et al. (2022)	CTC, AWC, CVC, Crying, APCU-C + V, electronic noise, vegetative sounds	—	Clinic	Pre-determined time	NR	Y
Oller et al. (2010)	NR	—	Home	All day	NR	N
Oller et al. (2015)	NR; A/A	—	Home	All day	NR	N
Pawar et al. (2017)	AWC, CVC, Noise, Silence	—	Home, Clinic	NR	Y	Y
Rankine et al. (2017)	CVC, Key Child	—	Home	All day	NR	Y
Sabatino-Devito et al. (2016)	CTC, AWC, CVC	—	Home	All day	15.95 h; range 4.8–16 h	Y

(Continued)

**Table 2.** (Continued)

Study	LENA Information	FAN-CHA	Home	All day	NR	N
Seidl et al. (2018)	CTC, CVC, MAN/FAN, AVA, CHA	FAN-CHA turns extracted with Praat	Home	Pre-determined time	3, 20-min recordings	N
Stolte (2017)	CTC, AWC, CVC	—	Clinic	Pre-determined time	87.97 (18.76) min	N
Sulek et al. (2022)	CTC, CVC	—	Home	All day	15.84 h avg across 2 days for all families ( $SD = .97$ h)	N
Swanson et al. (2018)	CTC, AWC, CVC	—	Home	All day	At least 8 h; 94.77% were full 16 h	N
Swanson et al. (2019)	CTC, AWC	—	Home	Pre-determined time	NR	N
Trembath et al. (2019)	CVC, speech versus nonspeech	—	Clinic	Pre-determined time	NR	N
Vivanti et al. (2019)	CVC	—	School	Pre-determined time	45 min	N
Warlambont et al. (2014)	AWC, CVC, speech versus nonspeech	Extant software to measure contingency	Home, clinic, school	Pre-determined time	12 h	N
Warren et al. (2010)	CTC, AWC, CVC	—	Home, school	All day	NR	N
Wojnaroski et al. (2017)	AVA	IVD, ACPU-C + V	Home	All day	NR	N
Xu et al. (2014)	—	ACPU	Home	All day	NR	N
Yoder et al. (2013)	—	Vocal age equivalence	Home	All day	NR	N

Note. Italicized rows indicate studies from the gray literature search. CTC = Conversational Turn Count; CVC = Child Vocalization Count; AWC = Adult Word Count; LLAS = LENA Language and Autism Screen; ACPU = Average Count per Utterance; RVC = Reciprocal Vocal Contingency; MAN = Male Adult Near; FAN = Female Adult Near; ACPU-C + V = Average Count per Utterance-Consonants and Vowels; AQV = Automated Quantity of Vocalizations; AVA = Automated Vocal Analysis Developmental Age Score; CHA = Near Child; IVD = Infraphonological Vocal Development Score. — indicates not applicable. NR indicates not reported.

**Table 3.** Strengths and limitations of LENA in included studies.

Study	Reported strengths of LENA	Reported limitations of LENA
Avendaño (2022)	–	–
Bak et al. (2019)	–	<ul style="list-style-type: none"> <li>- May not be a valid measure of language or social communication</li> <li>- Picks up words that are not directed toward the child</li> <li>- May not be valid for minimally verbal children</li> <li>- May not be reliable with minimally verbal children</li> </ul>
Bredin-Oja et al. (2018)	<ul style="list-style-type: none"> <li>- Using LENA in short bursts could be a useful metric to monitor progress in clinical settings</li> </ul>	<ul style="list-style-type: none"> <li>- Low correlation with child metrics of language</li> <li>- Doesn't code/count overlapping speech</li> <li>- Using a peak sampling method may not provide an accurate representation of the rest of the day</li> </ul>
Bruyneel et al. (2021)	–	–
Burgess et al. (2013)	–	–
Carr et al. (2014)	–	–
Dunst et al. (2013)	–	–
Dykstra et al. (2013)	–	–
Fasano et al. (2021)	<ul style="list-style-type: none"> <li>- Allows for quantification of rich information without relying on complex and lengthy coding</li> <li>- Good recording device</li> </ul>	<ul style="list-style-type: none"> <li>- LENA is not as accurate as human transcription</li> <li>- Doesn't code/count overlapping speech</li> <li>- Does not capture language or what is actually being spoken</li> <li>- Adults captured by LENA may have been speaking to other children</li> <li>- LENA algorithms are not always accurate and often confuse adult female speech as child speech</li> </ul>
Ferguson et al. (2020)	–	–
Fleurisaint et al. (2017)	–	–
Fragoso (2016)	–	–
Ghai et al. (2013)	–	–
Harbison et al. (2018)	<ul style="list-style-type: none"> <li>- There can be value in including objective measures collected in naturalistic environments</li> <li>- Useful for assessing the efficacy of an intervention</li> </ul>	<ul style="list-style-type: none"> <li>- Type and quality of adult verbalizations are not captured</li> <li>- Nonverbal communication behavior is not captured</li> <li>- LENA unintentionally picks up other adults who were not speaking to the child</li> <li>- Does not capture conversational turns between children</li> <li>- Not ideal for use with older children</li> <li>- Not accurate at distinguishing between older children and adults</li> <li>- Not ideal for use with minimally verbal children</li> <li>- Does not capture the quality of the child's language environment</li> <li>- Adults may be talking to other children</li> <li>- Does not capture echolalia or scripting</li> <li>- Does not differentiate between adults</li> <li>- Doesn't capture fine-grained linguistic measures or linguistic diversity</li> </ul>
Harvánek et al. (2022)	–	–
Irvin et al. (2013)	<ul style="list-style-type: none"> <li>- Good to use as an intervention tool to provide feedback</li> </ul>	–
Jones et al. (2019)	–	–
Little et al. (2019)	–	–
Markfeld, Feldman, Daly, et al. (2023)	<ul style="list-style-type: none"> <li>- Ecologically valid measure of infants' home language environments</li> <li>- Improves cost- and time-effectiveness of research</li> </ul>	<ul style="list-style-type: none"> <li>- LENA could be sensitive to change in preverbal children</li> <li>- Good for intervention measurement</li> </ul>
McDaniel (2017)	–	–
McDaniel, Yoder, Crandall, et al. (2020)	<ul style="list-style-type: none"> <li>- LENA may be restricted to children of a particular language level or across a certain time range</li> <li>- LENA may misclassify adult speakers with each other</li> <li>- No way to confirm whether an adult was speaking to a target child</li> </ul>	<ul style="list-style-type: none"> <li>(Continued)</li> </ul>

**Table 3. (Continued)**

Study	Reported strengths of LENA	Reported limitations of LENA
McDaniel, Yoder, Estes, & Rogers (2020)	- More affordable and less time-intensive than human coding	- LENA was not as predictive of later language skills as human coding was
Meera (2019)	-	- LENA has a high rate of misclassification of speakers
Moffitt et al. (2022)	- LENA measurements may reflect underlying autism-related characteristics and can inform clinicians early on	- LENA has a high rate of misclassification of speakers
Oller et al. (2010)	-	-
Oller et al. (2015)	- LENA may be a potential form of vocal assessment with no human intervention	-
Pawar et al. (2017)	-	- LENA mislabeled/misclassified some speakers
Rankine et al. (2017)	-	- LENA did not have high accuracy with child vocalization estimates
Sabatos-Derito (2016)	- Home-based computerized measurements are scalable, reliable, and quantitative outcome measure for use in clinical trials or treatments for social communication	- LENA is not appropriate for children with echolalia or other atypical nonspeech sounds
Seidl et al. (2018)	-	- Is not reliable with minimally verbal children
Stoltz (2017)	- Easy to use, access, and sync with video data - The hardware itself is very reliable for families and researchers	-
Sulek et al. (2022)	-	- Unclear whether LENA is a good predictor of later skills
Swanson et al. (2018)	- Automatic feedback from the LENA system helps Parents and caregivers provide a rich language environment	- Vocalization counts may not be valid, even when suggested age range is used
Swanson et al. (2019)	-	- Other people in the home may alter the counts
Trembath et al. (2019)	-	- LENA cannot be used in a wide variety of settings
Vivant et al. (2019)	-	-
Warlaumont et al. (2014)	-	- Not able to classify nondirected echolalia
Warren et al. (2010)	- Could be valuable for monitoring the effectiveness of an intervention and for monitoring language at home	-
Wojnaroski et al. (2017)	- LENA-derived indices are useful as predictors of future spoken vocabulary for preverbal autistic preschoolers	- Cannot measure important nonverbal measures
	- LENA is a cost-effective and time-efficient method of measuring children's vocal development in everyday practice	- All vocalizations are considered equally. A short, single-syllable vocalization is counted the same as a full sentence
Xu et al. (2014)	-	- Some measures derived by LENA are not available through the LENA software, so they're not accessible to all clinicians/researchers
Yoder et al. (2013)	-	-

Note: Italicized rows indicate studies from the gray literature search. LENA = Language ENvironment Analysis System, CVC = Child Vocalization Count. – indicates no strengths and/or limitations reported.

accuracy of LENA in identifying and labeling the child language environment, and human transcription/coding was central to their methodology. Researchers in three other studies (Bak et al., 2019; Fasano et al., 2021; Moffitt et al., 2022) manually coded a small portion of their audio segments for quality purposes. Bak et al. (2019) manually coded one-hour samples from their recordings, while Fasano et al. (2021) and Moffitt et al. (2022) coded a selection of LENA-identified vocalizations. Methods for determining agreement between human methods and LENA-derived categorizations varied widely, as did the resulting reports of agreement (agreement between human coded and LENA variables ranged from 23% to 87% across studies and variables coded).

**Other measures used alongside LENA.** We extracted whether studies used additional measures to capture aspects of language, child behavior, and/or the environment more broadly (e.g. GPS technology). There was a wide range in additional measures used, with 29 different measures used across studies. Thirty studies used standardized (e.g. Mullen Scales of Early Learning) or naturalistic (e.g. communication sample) assessments in tandem with or in comparison to LENA to measure speech and/or language abilities. Eleven of the 42 studies did not use any other measures in addition to LENA. The range of additional measures used across studies was zero to eight, and the average number of additional measures used was 1.95.

### ***Qualitatively coded strengths for LENA use in autism research***

Content analysis of explicitly-stated strengths and weaknesses noted in the included studies resulted in the following themes addressing the *strengths* of LENA use in autism research. These themes are summarized in Table 3.

**LENA is ecologically valid.** A strength highlighted in two studies is that LENA is an objective and ecologically valid measure of speech (Hardan, 2016; Markfeld, Feldman, Bordman, et al., 2023).

**LENA provides quick feedback to caregivers and clinicians.** Researchers in five studies suggested that with its quick processing and user-friendly output, LENA is a useful tool for giving real-time feedback to give clinicians and caregivers (Irvin et al., 2013; Swanson et al., 2018; Warren et al., 2010; Woynaroski et al., 2017). Automated LENA results may be used in everyday practice to monitor clinical progress and to inform caregiver speech (Bredin-Oja et al., 2018; Swanson et al., 2018; Woynaroski et al., 2017).

**LENA is a useful outcome measure in intervention research.** LENA was noted as a strong asset to intervention-based research in six studies, with many researchers citing the quantifiable measures of naturalistic speech as a measure of speech and language intervention outcomes (Bredin-Oja et al., 2018; Hardan, 2016; Irvin et al., 2013; McDaniel, Yoder, Crandall, et al., 2020; Sabatos-DeVito, 2016; Warren et al., 2010).

**LENA is an informative tool for monitoring speech and autism characteristics.** Researchers in four studies suggested that LENA has potential use in informing clinicians of early autism characteristics in child speech (Moffitt et al., 2022; Oller, 2015) and in monitoring speech development in preverbal autistic children (McDaniel, 2017; Woynaroski et al., 2017).

**LENA is time- and cost-effective.** Researchers in four studies noted that the LENA system is more cost-effective and less time-intensive than manual transcription and coding, providing users with an array of informative measurements without the hours of labor that would be otherwise required to achieve the same results (Fasano et al., 2021; Markfeld, Feldman, Bordman, et al., 2023; McDaniel, Yoder, Estes, & Rogers, 2020; Woynaroski et al., 2017).

**LENA is user-friendly.** Stolte (2017) noted the accessibility of the software and ease of syncing the data with other software. Not only does the system make measurement of a child's home language environment easier, but researchers in two studies noted the ease of use for families as well: The recording device is durable with a long battery life, allowing for use throughout daily activities in multiple settings (Ferguson et al., 2020; Stolte, 2017).

### ***Coded limitations and recommendations for LENA use in autism research***

Thematic analysis of the included studies resulted in the following themes addressing the *limitations* of LENA use as explicitly stated by included study authors, as well as recommendations made by those authors to address such limitations:

**LENA commonly misclassifies speakers.** Among the most commonly expressed limitation of the LENA system was the frequent misclassification of child and adult speakers, resulting in validity concerns for eight studies (Bak et al., 2019; Fasano et al., 2021; Ferguson et al., 2020; McDaniel, Yoder, Crandall, et al., 2020; Moffitt et al., 2022; Pawar et al., 2017; Rankine et al., 2017; Swanson et al., 2018). This concern was cited in toddler-based studies as well as in studies with participants outside the recommended age range of LENA use. Such concerns led to calls for future

work to examine the validity of these automated measures (Bak et al., 2019; Bredin-Oja et al., 2018; Swanson et al., 2018; Woynaroski et al., 2017). Researchers recommended using the LENA software either within the recommended age range or alongside human annotation to avoid misclassification errors (Ferguson et al., 2020; Jones et al., 2019; Oller et al., 2010; Pawar et al., 2017; Sulek et al., 2022). Others suggested using data from longer recordings or from multi-day recordings to increase the stability of the estimates (Bak et al., 2019; Harbison et al., 2018; McDaniel, Yoder, Crandall, et al., 2020; Sulek et al., 2022; Trembath et al., 2019; Woynaroski et al., 2017).

**LENA does not capture the nature or complexity of child speech.** Two main concerns regarding the complexity of child speech arose from nine studies. First, while the system provides users with a quantified metric of speech, LENA does not capture the *intentionality* of child vocalizations (Avendaño, 2022; Bredin-Oja et al., 2018; Dykstra et al., 2013), or the type and quality of the adult vocalizations directed toward children (Fragoso, 2016; Irvin et al., 2013; Little et al., 2019). Researchers using the software are restricted to viewing frequency data alone without any context (Stolte, 2017). Second, other researchers identified the software's dichotomous coding (present or not) of child vocalizations as a limitation, as it does not allow for insights into fine-grained language abilities (Markfeld, Feldman, Bordman, et al., 2023; Warren et al., 2010). To address this limitation, researchers suggested the use of LENA in tandem with human transcriptions as well as formal language measures (Bak et al., 2019; Carr et al., 2014; Hardan, 2016; Irvin et al., 2013; Rankine et al., 2017; Trembath et al., 2019).

**LENA data does not consider echolalia.** The lack of information surrounding the content or complexity of speech from children/adults led researchers in seven studies to warn against the use of LENA variables when working with children with high levels of echolalia in their speech (Dykstra et al., 2013; Little et al., 2019; Rankine et al., 2017; Trembath et al., 2019). Indeed, the use of echolalia or scripting by children was thought to artificially inflate reported CVC numbers (Dykstra et al., 2013; Trembath et al., 2019). In addition, multiple researchers recommended not using LENA variables with minimally verbal autistic children (Bak et al., 2019; Bredin-Oja et al., 2018; Jones et al., 2019; Rankine et al., 2017).

**LENA does not capture the holistic language environment.** A limitation of LENA, as highlighted by five studies, is that it cannot capture overlapping speech or conversational turns between children (Bruyneel et al., 2021; Ferguson et al., 2020; Irvin et al., 2013). Researchers urged the consideration of the recording setting in analysis, particularly when using LENA in school settings or other environments with multiple children/adults (Bruyneel et al., 2021;

Dykstra et al., 2013; Ferguson et al., 2020). Other important aspects of communication in recording environments are not captured by LENA, such as nonverbal communication or the visual, spatial, and sensory environments (Dykstra et al., 2013; Irvin et al., 2013; Warren et al., 2010). This limitation was overcome by the studies that synced LENA data with other devices that recorded the child's location or the child's visual environment (Fasano et al., 2021; Little et al., 2019).

**LENA does not distinguish between adults and picks up speech not directed toward the target child.** Concerns surrounding the interpretability of the Adult Word Count variable were also expressed in eight studies, as LENA software has no way of distinguishing between adults in the home. In environments with multiple children or multiple adults, the LENA system classifies all nearby adult vocalizations when computing Adult Word Count, leaving an unclear idea of which of those vocalizations were directed toward the target child (Bak et al., 2019; Dykstra et al., 2013; Harbison et al., 2018; Irvin et al., 2013; Little et al., 2019; Markfeld, Feldman, Bordman, et al., 2023; Swanson et al., 2018). Moreover, for intervention studies involving caregiver training, it is unclear whether adult speech captured by LENA came from the caregiver undergoing the training, or another adult (McDaniel, Yoder, Crandall, et al., 2020).

## Discussion

The purpose of this systematic review was to determine how LENA is being used in autism research. By conducting an in-depth examination into the strategies reported by researchers, we summarized the methods employed with LENA in this field as well as the highlighted strengths and limitations.

### How is LENA being used in autism research?

Most autism research employed LENA for exploratory purposes, using LENA recordings and resulting metrics to measure language input and development and to examine autism-specific relationships to LENA metrics. Multiple researchers called for research around additional norming and validation of LENA (Bak et al., 2019; Dykstra et al., 2013; Fragoso, 2016; Markfeld, Feldman, Bordman, et al., 2023; Swanson et al., 2018), which was the second most popular purpose of autism research using LENA. While relatively few studies used LENA in intervention settings, those that did identified the LENA system as a strong asset to their intervention protocols. Despite the promotion of LENA use in the early detection and screening for autism (e.g. Richards et al., 2010), only two of the included studies (Carr et al., 2014; Oller et al., 2010) used LENA for this purpose. While the lack of supporting evidence does not

mean that LENA metrics are not useful for discriminating between autistic and non-autistic children, it does suggest that fewer studies have focused on examining LENA's potential for early detection than expected. This may reflect a hesitancy toward using autism screeners by clinicians (Watson & Crais, 2013) as well as a shift in research priorities in recent years away from identifying causes for autism and toward services and supports for diagnosed individuals (Fletcher-Watson et al., 2017; Gotham et al., 2015; Pellicano et al., 2014).

The majority of included studies had participant samples comprised of autistic children; the remaining studies largely focused on high-likelihood siblings. Very few comparisons between the speech of autistic and non-autistic children were made. Instead, many researchers used LENA to compare the language environments of one child across multiple settings, timepoints, or forms of measurement. These study aims are largely in line with the goal within autism research to have a more ecologically valid understanding of speech and language for autistic children (Bergelson et al., 2019; Woynaroski et al., 2017) and also suggest that autism researchers may be using LENA in similar ways that research with non-autistic children is conducted (Cychosz & Cristia, 2021). More support for this consistency comes from the similar use of LENA variables across studies of autism and neurotypical language development (i.e. CVC, AWC, CTC; Cristia et al., 2020).

Most studies using LENA for autism research did so with samples entirely or partially within the age window LENA variables were normed on (2–48 months). Although a stable autism diagnosis can be obtained as early as 18 months (Hyman et al., 2020; Zwaigenbaum et al., 2016), the average age of diagnosis remains at around four years old or older (Shattuck et al., 2009; Zuckerman et al., 2017), placing the average autistic child at the time of diagnosis outside the recommended age window for analysis using LENA metrics. Future research should aim to extend the benefits of daylong recordings to older children, enabling researchers and clinicians to capture ecologically valid indices of speech without the resources required for human coding.

LENA recordings were most commonly collected in the homes of participating families. This is unsurprising given LENA guidelines to collect data in the home for validity purposes, but, as emphasized by multiple authors of the included studies (Burgess et al., 2013; Dykstra et al., 2013; Ferguson et al., 2020; Little et al., 2019), understanding speech and language of autistic children in school and community settings is critical. For this reason, we join these authors in calling for the continued development of naturalistic and efficient measures of speech and language in multiple settings where the automated LENA software is more prone to speaker misclassification and overidentification of child/adult vocalizations. Much of this work is complicated by the ethical barriers to collecting audio recordings in community settings. A recent article suggests

that in addition to following regional third-party consent laws, researchers may provide participants with information cards or short verbal explanations to give to those who may be in range of the recording device, such as visitors to a home, nearby families in public spaces, or classroom staff (Cychosz et al., 2020).

### ***Considerations and recommendations for researchers and clinicians***

***Compatibility with target child and setting.*** When deciding whether to use LENA, researchers and clinicians should first consider that compatibility may vary depending on child characteristics. LENA may not be the best fit for all children, and a given child's age (e.g. is LENA is valid in this age range?), sensory profile (e.g. will the child be comfortable wearing a LENA device for lengthy recording periods?), and frequency of echolalia and/or repetitive speech (e.g. would vocalization counts be higher than expected?) should be considered before use. The anticipated recording environment should also be considered: Automated indices of speech from recordings in settings with multiple adults or children may be less accurate.

As noted in the stated limitations of LENA in the included studies, it is important to consider the large limitation that LENA doesn't capture the intentionality or content of speech. Specifically, when considering this limitation within the context of speech and language profiles in autism, large amounts of echolalia and/or repetitive speech could inflate CVC. As such, we do not recommend using LENA's automated child vocalization counts as a primary measure of speech production for autistic children who are reported to produce frequent repetitive speech. To overcome some of these limitations, users may want to integrate data checks such as human transcription or annotation into their methods, especially in samples where there are great individual differences in speech production due to autism features such as echolalia. In addition, users should consider pairing the LENA recording device with alternative processing methods. This is particularly important when studying the language environment of autistic children who are outside the recommended age range of 2–48 months, but may reduce the possibility for speaker misclassification, which was reported in studies with participants both within and outside this window. There are multiple extant automated processing pipelines commonly used in the field, including software from the ChildProject (Cristia et al., 2023; Gautheron et al., 2023), the Automatic Linguistic Unit Count Estimator (ALICE; Räsänen et al., 2021), and a recent algorithm developed by Bang et al. (2022) to classify target child-directed speech and times when the target child is awake versus sleeping. These open-source extant processing methods may overcome age-related limitations of LENA software.

**Capturing language.** A key limitation noted by multiple studies is that LENA cannot provide all information needed to capture language ability in autistic children. To enrich the information provided by the LENA vocalization counts, we recommend using an additional measure of language or development (Adams, 2002; Tager-Flusberg et al., 2009) such as the MacArthur-Bates Communicative Developmental Inventories (Fenson et al., 2007) or the Preschool Language Scales, 5th Edition (Zimmerman et al., 2011). Using multiple measurements of language may better capture the heterogeneity of language in autism.

**Reliability and validity.** Not all LENA indices have been validated within the autistic population and some may vary in stability. Although some LENA variables have been validated against human transcription in non-autistic populations (Cristia et al., 2020; Ganek & Eriks-Brophy, 2018; Levin-Asher et al., 2023), the evidence on the validity of LENA variables in autistic and high likelihood children is sparse. One factor that may limit the validity of LENA indices is the stability of these variables. Results from four empirical studies may guide autism researchers on how many daylong recordings are needed to obtain stable estimates of LENA variables, specifically with infant siblings and autistic preschoolers (i.e. Harbison et al., 2018; Markfeld, Feldman, Bordman, et al., 2023; Woynaroski et al., 2017; Yoder et al., 2013). These papers found that multiple daylong recordings are needed to obtain stable estimates of LENA variables, and that the number of recording days recommended varies according to the variable of interest. If researchers are interested in utilizing LENA in other age ranges on the spectrum or in other clinical phenotypes (e.g. school-aged children, adolescents with limited spoken language), we recommend that they first evaluate the stability of these LENA indices in their sample by conducting a Generalizability and Decision study (Cronbach's et al., 1963). Future data on the stability of LENA variables will allow the field to set more consistent standards regarding how many recording days should be collected.

It should also be noted that stability alone is not sufficient to assume validity of the variables of interest (Yoder et al., 2018). Therefore, stable estimates of variables should not be considered sufficient evidence that these metrics are valid, particularly when other factors (e.g. sleep/wake time, LENA software errors in speaker classification) may be inconsistent and could influence measurement error unevenly across participants. For example, rates of LENA indices (e.g. AWC, CVC) could be underestimated if infant sleeping hours are not removed from recordings prior to the calculation of average counts. Current approaches to account for sleep/wake time include trimming all recordings so they begin after the child wakes and end before the child is sleeping (e.g. Warlaumont et al., 2014), or to individually trim recordings based on when a child wakes up and goes to sleep so that rates are

not impacted by sleeping times (e.g. Swanson et al., 2019). Further research will be needed to consider these types of measurement error (see Bang et al., 2022 for potential methodological advances).

**Reporting standards for LENA research.** As research using daylong, naturalistic recordings in clinically relevant populations progresses, users can take several steps to ensure that their work can move the field forward. When publishing work in this area, we encourage researchers to share specific information regarding the sex, race, home language, and age (average and range) of participants in their study. Age is a critical factor in evaluating the decision to use LENA-measured indices of speech for analysis and can inform the decisions of others aiming to conduct a similar analysis. In addition, fewer than half of the studies included in this review reported the race of their participants, a factor that can provide important cultural context regarding the reproducibility of this work and the possibility that these findings could generalize to all individuals on the autism spectrum (e.g. Millager et al., 2024). A crucial first step to increasing diversity and representation in autism research is to report participant race, and to consider how sociocultural factors may impact findings (Steinbrenner et al., 2022). Further, we encourage researchers to report details regarding their recording and analysis of LENA variables. Although the LENA guidelines encourage recording in the home and for a 16 hour period, these conditions cannot be assumed if not reported. Data transformation practices (e.g. converting raw values to averages, trimming recordings) contributes to the validity of the data, and these practices should be shared to allow for comparison across studies as well as to allow future LENA users to select the most appropriate method of data management.

### **Strengths and limitations of this review**

A significant strength of this systematic review is the inclusion of gray literature and articles published in languages other than English, thus widening the available information regarding use of the LENA system in autism research to date (Konno et al., 2020; Paez, 2017). To our knowledge, this is the most exhaustive list of published studies using LENA; however, there may have been studies not detected using our search terms. In addition, due to the inconsistency of study design, data collection, and reporting across the included studies, we were unable to meta-analyze results of interest such as language growth or outcomes. One potential solution and future direction is for authors to share anonymized .its files to enable further integration of findings across studies. Finally, our qualitative coding of the strengths and limitations of LENA as stated by researchers in the included studies was limited to those that were explicitly provided in the study texts. The

LENA technology itself was not mentioned in the discussion of the majority of the included studies, as it is not required in standard manuscript guidelines. As such, it is unclear whether authors that did *not* explicitly describe a given strength or limitation of LENA would also share these views expressed herein.

### Future directions

As stated above, additional work is needed to assess the stability of LENA indices in autism. In addition, researchers should continue to examine how LENA indices are associated with language abilities. Conclusions about these relations have varied widely in existing work both within and outside of autism research (Cychosz & Cristia, 2021), and may be particularly complex for autism. Future research should continue to develop ways to measure language use in older autistic individuals. Finally, when publishing data using LENA, authors should provide detailed information about their methods to increase generalizability and transparency.

### Conclusion

This systematic review is, to our knowledge, the first to synthesize methods for using the LENA system in the field of autism research. We found that autism researchers have used LENA for a variety of purposes, but largely to characterize the speech environment of autistic children or children with higher likelihood of autism. For the most part, studies used LENA within the recommended age and recording length/location guidelines. Described strengths of the LENA system included its ecological validity, cost/time-effectiveness, and user accessibility. Described limitations included high rates of speaker misclassification and lack of information surrounding language context/content. This review provides a useful starting point for researchers and clinicians who are using LENA with autistic children, as well as considerations for future directions for those who use LENA in research.

### Acknowledgements

The authors thank Tiffany Woynaroski and members of the Daylong Audio Recordings of Children's Linguistic Environments (DARCLE) group for providing feedback on versions of this manuscript.

### Author contributions

O.C.P. and C.H. initiated the idea for this review with consultation from S.T.W. O.C.P. and J.G. completed the PROSPERO registration with supervision from C.H. S.T.W. completed the search strategy and provided the screening infrastructure via Covidence. O.C.P., J.G., and C.H. completed the abstract screening, and O.C.P., J.E.M., and C.H. completed the full-text screening. O.C.P., J.E.M., and M.S. conceptualized the coding scheme.

O.C.P., J.E.M., M.K., A.J.N., and J.G. completed the article coding and quality assessment. O.C.P. and J.E.M. drafted the manuscript with edits and contributions from all other authors. All authors read and approved the final manuscript.

### Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: J.E.M., J.I.F., and M.S. have published work using LENA in autistic populations that were included in this review. They were not involved in the coding or synthesis of results pertaining to their own papers. The remaining authors have no conflicts of interests to disclose.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by training grants from the US Department of Education (H325D180099; O.C.P.) and the National Center for Advancing Translational Sciences (TL1 TR002244; J.I.F.), by a National Institute on Deafness and Other Communication Disorders grant (R01DC020186; J.E.M.), and from a predoctoral fellowship from the Autism Science Foundation (23-003; O.C.P.). Prior to submission for publication, a version of this research was presented as an oral presentation by O.C.P. and J.E.M. at the annual meeting for the American Speech and Hearing Association (ASHA) in 2022.

### ORCID iDs

Orla C. Putnam  <https://orcid.org/0000-0002-2215-0021>  
 Jennifer E. Markfeld  <https://orcid.org/0000-0002-7995-6077>  
 Sarah Towner Wright  <https://orcid.org/0000-0003-2076-8343>  
 Jacob I. Feldman  <https://orcid.org/0000-0002-5723-5834>  
 Jessica Goldblum  <https://orcid.org/0000-0003-4769-3935>  
 Clare Harrop  <https://orcid.org/0000-0003-3381-3473>

### Supplemental Material

Supplemental material for this article is available online.

### References

Adams, C. (2002). Practitioner review: The assessment of language pragmatics. *Journal of Child Psychology and Psychiatry*, 43(8), 973–987. <https://doi.org/10.1111/1469-7610.00226>

American Speech-Language-Hearing Association. (n.d.). *Autism (Autism Spectrum Disorder)*. <https://www.asha.org/public/speech/disorders/autism/>

Avendaño, S. M. (2022). *Language matters: LENA technology in research and practice*. Michigan State University.

Bak, M. Y. S., Plavnick, J. B., & Byrne, S. M. (2019). Vocalizations of minimally verbal children with autism spectrum disorder across the school year. *Autism*, 23(2), 371–382.

Bang, J. Y., Kachergis, G., Weisleder, A., & Marchman, V. A. (2022). *Evaluating the feasibility of an automated classifier for target-child-directed speech from LENA recordings*. PsyArXiv. <https://doi.org/10.31234/osf.io/5g2jb>

Barokova, M., & Tager-Flusberg, H. (2020). Commentary: Measuring language change through natural language samples. *Journal of Autism and Developmental Disorders*, 50(7), 2287–2306. <https://doi.org/10.1007/s10803-018-3628-4>

Benevides, T. W., Shore, S. M., Palmer, K., Duncan, P., Plank, A., Andresen, M.-L., Caplan, R., Cook, B., Gassner, D., Hector, B. L., Morgan, L., Nebeker, L., Purkis, Y., Rankowski, B., Wittig, K., & Coughlin, S. S. (2020). Listening to the autistic voice: Mental health priorities to guide research and practice in autism from a stakeholder-driven project. *Autism*, 24(4), 822–833. <https://doi.org/10.1177/1362361320908410>

Bergelson, E., Amatuni, A., Dailey, S., Koorathota, S., & Tor, S. (2019). Day by day, hour by hour: Naturalistic language input to infants. *Developmental Science*, 22(1), e12715. <https://doi.org/10.1111/desc.12715>

Billstedt, E., Gillberg, C., & Gillberg, C. (2005). Autism after adolescence: Population-based 13- to 22-year follow-up study of 120 individuals with autism diagnosed in childhood. *Journal of Autism and Developmental Disorders*, 35(3), 351–360. <https://doi.org/10.1007/s10803-005-3302-5>

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>

Bredin-Oja, S. L., Fielding, H., Fleming, K. K., & Warren, S. F. (2018). Clinician vs. machine: Estimating vocalizations rates in young children with developmental disorders. *American Journal of Speech-Language Pathology*, 27(3), 1066–1072. [https://doi.org/10.1044/2018\\_AJSLP-17-0016](https://doi.org/10.1044/2018_AJSLP-17-0016)

Bruyneel, E., Demurie, E., Boterberg, S., Warreyn, P., & Roeyers, H. (2021). Validation of the Language ENvironment Analysis (LENA) system for Dutch. *Journal of Child Language*, 48(4), 765–791. <https://doi.org/10.1017/S0305000920000525>

Burgess, S., Audet, L., & Harjusola-Webb, S. (2013). Quantitative and qualitative characteristics of the school and home language environments of preschool-aged children with ASD. *Journal of Communication Disorders*, 46(5–6), 428–439. <https://doi.org/10.1016/j.jcomdis.2013.09.003>

Butler, L. K., Kiran, S., & Tager-Flusberg, H. (2020). Functional near-infrared spectroscopy in the study of speech and language impairment across the life span: A systematic review. *American Journal of Speech-Language Pathology*, 29(3), 1674–1701. [https://doi.org/10.1044/2020\\_AJSLP-19-00050](https://doi.org/10.1044/2020_AJSLP-19-00050)

Carr, J., Xu, D., & Yoshinaga-Itano, C. (2014). Language ENvironment analysis language and autism screen and the child development inventory social subscale as a possible autism screen for children who are deaf or hard of hearing. *Seminars in Speech and Language*, 35(04), 266–275. <https://doi.org/10.1055/s-0034-1389099>

Channell, M. M., Loveall, S. J., Conners, F. A., Harvey, D. J., & Abbeduto, L. (2018). Narrative language sampling in typical development: Implications for clinical trials. *American Journal of Speech-Language Pathology*, 27(1), 123–135. [https://doi.org/10.1044/2017\\_AJSLP-17-0046](https://doi.org/10.1044/2017_AJSLP-17-0046)

Cristia, A., Gautheron, L., & Colleran, H. (2023). Vocal input and output among infants in a multilingual context: Evidence from long-form recordings in Vanuatu. *Developmental Science*, 26(4), Article e13375. <https://doi.org/10.1111/desc.13375>

Cristia, A., Lavechin, M., Scaff, C., Soderstrom, M., Rowland, C., Räsänen, O., Bunce, J., & Bergelson, E. (2020). A thorough evaluation of the Language Environment Analysis (LENA) system. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-020-01393-5>

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology*, 16(2), 137–163. <https://doi.org/10.1111/j.2044-8317.1963.tb00206.x>

Cychosz, M., & Cristia, A. (2021). *Using big data from long-form recordings to study development and optimize societal impact*. <https://www.sciencedirect.com/science/article/abs/pii/S0065240721000434>

Cychosz, M., Romeo, R., Soderstrom, M., Scaff, C., Ganek, H., Cristia, A., Casillas, M., De Barbaro, K., Bang, J. Y., & Weisleder, A. (2020). Longform recordings of everyday life: Ethics for best practices. *Behavior Research Methods*, 52, 1951–1969.

Dillon, E. F., Kanne, S., Landa, R. J., Annett, R., Bernier, R., Bradley, C., Carpenter, L., Kim, S. H., Parish-Morris, J., Schultz, R., & Wodka, E. L., The SPARK Consortium, & Wodka, E. L. (2021). Sex differences in autism: Examining intrinsic and extrinsic factors in children and adolescents enrolled in a national ASD cohort. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-021-05385-y>

Dunst, C. J., Trivette, C. M., Hamby, D. W., Prior, J., & Derryberry, G. (2013). Effects of Child-Robot interactions on the vocalization production of young children with disabilities. *Social Robots Research Reports*, 4, 1–10.

Dykstra, J. R., Sabatos-DeVito, M. G., Irvin, D. W., Boyd, B. A., Hume, K. A., & Odom, S. L. (2013). Using the Language Environment Analysis (LENA) system in preschool classrooms with children with autism spectrum disorders. *Autism*, 17(5), 582–594. <https://doi.org/10.1177/1362361312446206>

Eaves, L. C., & Ho, H. H. (2008). Young adult outcome of autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 38(4), 739–747. <https://doi.org/10.1007/s10803-007-0441-x>

Edmunds, S. R., Kover, S. T., & Stone, W. L. (2019). The relation between parent verbal responsiveness and child communication in young children with or at risk for autism spectrum disorder: A systematic review and meta-analysis. *Autism Research*, 12(5), 715–731. <https://doi.org/10.1002/aur.2100>

Fasano, R. M., Perry, L. K., Zhang, Y., Vitale, L., Wang, J., Song, C., & Messinger, D. S. (2021). A granular perspective on inclusion: Objectively measured interactions of preschoolers with and without autism. *Autism Research*, 14(8), 1658–1669. <https://doi.org/10.1002/aur.2526>

Fenson, L., Marchman, V., Thal, D., Dale, P., Reznick, S., & Bates, E. (2007). *The MacArthur-Bates Communicative Development Inventories*. Brookes Publishing. <https://products.brookespublishing.com/The-MacArthur-Bates-Communicative-Development-Inventories-Users-Guide-and-Technical-Manual-Second-Edition-P78.aspx>

Ferguson, E. F., Nahmias, A. S., Crabbe, S., Liu, T., Mandell, D. S., & Parish-Morris, J. (2020). Social language opportunities

for preschoolers with autism: Insights from audio recordings in urban classrooms. *Autism*, 24(5), 1232–1245. <https://doi.org/10.1177/1362361319894835>

Fletcher-Watson, S., Apicella, F., Auyeung, B., Beranova, S., Bonnet-Brilhault, F., Canal-Bedia, R., Charman, T., Chericoni, N., Conceição, I. C., Davies, K., Farroni, T., Gomot, M., Jones, E., Kaale, A., Kapica, K., Kawa, R., Kylliäinen, A., Larsen, K., Lefort-Besnard, J., & Yirmiya, N. (2017). Attitudes of the autism community to early autism research. *Autism*, 21(1), 61–74. <https://doi.org/10.1177/1362361315626577>

Fleurissant, R. M. (2017). *Early developmental trajectories of social contingency predict language outcome in toddlers with ASD*. International Society for Autism Research.

Fragoso, N. C. (2016). *Language usage and exposure by 5- and 6-year-old children with autism spectrum disorder and their families living in the imperial valley of California*. Claremont Graduate University and San Diego State University.

Ganek, H. V., & Eriks-Brophy, A. (2018). A concise protocol for the validation of Language ENvironment Analysis (LENA) conversational turn counts in Vietnamese. *Communication Disorders Quarterly*, 39(2), 371–380. <https://doi.org/10.1177/1525740117705094>

Gautheron, L., Rochat, N., & Cristia, A. (2023). Managing, storing, and sharing long-form recordings and their annotations. *Language Resources and Evaluation*, 57(1), 343–375. <https://doi.org/10.1007/s10579-022-09579-3>

Georgiou, N., & Spanoudis, G. (2021). Developmental language disorder and autism: Commonalities and differences on language. *Brain Sciences*, 11(5), 589. <https://doi.org/10.3390/brainsci11050589>

Ghai, S. (2013, May 2). *Developmental profiling of voice quality in infants at risk of autism*. International Society for Autism Research.

Gilkerson, J., & Richards, J. (2020). *A guide to understanding the design and purpose of the LENA system* (Technical Report LTR-12; LENA Technical Report). LENA Foundation. [https://www.lenafoundation.org/wp-content/uploads/2020/07/LTR-12\\_How\\_LENA\\_Works.pdf](https://www.lenafoundation.org/wp-content/uploads/2020/07/LTR-12_How_LENA_Works.pdf)

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Kimbrough, O. 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](https://doi.org/10.1044/2016_AJSLP-15-0169)

Gotham, K., Marvin, A. R., Taylor, J. L., Warren, Z., Anderson, C. M., Law, P. A., Law, J. K., & Lipkin, P. H. (2015). Characterizing the daily life, needs, and priorities of adults with autism spectrum disorder from Interactive Autism Network data. *Autism*, 19(7), 794–804. <https://doi.org/10.1177/1362361315583818>

Greenwood, C. R., Schnitz, A. G., Irvin, D., Tsai, S. F., & Carta, J. J. (2018). Automated language environment analysis: A research synthesis. *American Journal of Speech-Language Pathology*, 27(2), 853–867. [https://doi.org/10.1044/2017\\_AJSLP-17-0033](https://doi.org/10.1044/2017_AJSLP-17-0033)

Harbison, A. L., Woynaroski, T. G., Tapp, J., Wade, J. W., Warlaumont, A. S., & Yoder, P. J. (2018). A new measure of child vocal reciprocity in children with autism spectrum disorder: Child vocal reciprocity. *Autism Research*, 11(6), 903–915. <https://doi.org/10.1002/aur.1942>

Hardan, A. Y. (2016). *The LENA system in clinical trials: Evidence from pivotal response treatment studies*. Journal of the American Academy of Child & Adolescent Psychiatry.

Harrop, C., Libsack, E., Bernier, R., Dapretto, M., Jack, A., McPartland, J. C., Van Horn, J. D., Webb, S. J., & Pelphrey, K. & the GENDAAR Consortium. (2021). Do biological sex and early developmental milestones predict the age of first concerns and eventual diagnosis in Autism Spectrum Disorder? *Autism Research*, 14(1), 156–168. <https://doi.org/10.1002/aur.2446>

Harvánek, R., Kučera, M., Du, J., Li, J., & Kantor, J. (2022). The effect of musical sensorimotor integrative therapy on the speech of a child with autism. *Rehabilitace a Fyzikální Lékařství*, 29(4), 194–202. <https://doi.org/10.48095/crrhf2022194>

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O’Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, 34(4), 285–291. <https://doi.org/10.3233/EFI-180221>

Hyman, S. L., Levy, S. E., & Myers, S. M. Council on Children with Disabilities Section on Developmental and Behavioral Pediatrics Kuo, D. Z., Apkon, S., Davidson, L. F., Ellerbeck, K. A., Foster, J. E. A., Noritz, G. H., Leppert, M. O., Saunders, B. S., Stille, C., Yin, L., Weitzman, C. C., Childers, D. O., Levine, J. M., Peralta-Carcelen, A. M., Poon, J. K., & Bridgemohan, C. (2020). Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics*, 145(1), Article e20193447. <https://doi.org/10.1542/peds.2019-3447>

Irvin, D. W., Hume, K., Boyd, B. A., McBee, M. T., & Odom, S. L. (2013). Child and classroom characteristics associated with the adult language provided to preschoolers with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 7(8), 947–955. <https://doi.org/10.1016/j.rasd.2013.04.004>

Jones, R. M., Plesa Skwerer, D., Pawar, R., Hamo, A., Carberry, C., Ajodan, E. L., Caulley, D., Silverman, M. R., McAdoo, S., Meyer, S., Yoder, A., Clements, M., Lord, C., & Tager-Flusberg, H. (2019). How effective is LENA in detecting speech vocalizations and language produced by children and adolescents with ASD in different contexts? *Autism Research*, 12(4), 628–635. <https://doi.org/10.1002/aur.2071>

Konno, K., Akasaka, M., Koshida, C., Katayama, N., Osada, N., Spake, R., & Amano, T. (2020). Ignoring non-English-language studies may bias ecological meta-analyses. *Ecology and Evolution*, 10(13), 6373–6384. <https://doi.org/10.1002/ece3.6368>

La Valle, C., Shen, L., Shih, W., Kasari, C., Lord, C., & Tager-Flusberg, H. (2024). Real-time coded measures in natural language samples capture change over time in minimally verbal autistic children. *Autism Research*, 17(6), 1287–1293. <https://doi.org/10.1002/aur.3142>

Levin-Asher, B., Segal, O., & Kishon-Rabin, L. (2023). The validity of LENA technology for assessing the linguistic environment and interactions of infants learning Hebrew

and Arabic. *Behavior Research Methods*, 55(3), 1480–1495. <https://doi.org/10.3758/s13428-022-01874-9>

Little, L. M., Rojas, J. P., Bard, A., Luo, Y., Irvin, D., & Rous, B. (2019). Automated measures to understand communication opportunities for young children with autism in the community: A pilot study. *OTJR: Occupation, Participation and Health*, 39(2), 124–130. <https://doi.org/10.1177/1539449219834911>

Markfeld, J. E., Feldman, J. I., Bordman, S. L., Daly, C., Santapuram, P., Humphreys, K. L., Keçeli-Kaysılı, B., & Woynaroski, T. G. (2023). Associations between caregiver stress and language outcomes in infants with autistic and non-autistic siblings: An exploratory study. *Journal of Speech, Language, and Hearing Research*, 66(1), 190–205. [https://doi.org/10.1044/2022\\_JSLHR-22-00154](https://doi.org/10.1044/2022_JSLHR-22-00154)

Markfeld, J. E., Feldman, J. I., Daly, C., Santapuram, P., Bowman, S. M., Dunham-Carr, K., Suzman, E., Keçeli-Kaysılı, B., & Woynaroski, T. G. (2023). The stability and validity of automated indices of vocal development in infants with autistic and non-autistic siblings. *Journal of Speech, Language, and Hearing Research*, 66(12), 4934–4948. [https://doi.org/10.1044/2023\\_JSLHR-23-00069](https://doi.org/10.1044/2023_JSLHR-23-00069)

Matheis, M., Matson, J. L., Burns, C. O., Jiang, X., Peters, W. J., Moore, M., de Back, K. A., & Estabillo, J. (2017). Factors related to parental age of first concern in toddlers with autism spectrum disorder. *Developmental Neurorehabilitation*, 20(4), 228–235. <https://doi.org/10.1080/17518423.2016.1211186>

McCormick, C. E. B., Kavanagh, B. C., Sipsock, D., Righi, G., Oberman, L. M., Moreno De Luca, D., Gamsiz Uzun, E. D., Best, C. R., Jerskey, B. A., Quinn, J. G., Jewel, S. B., Wu, P., McLean, R. L., Levine, T. P., Tokadjian, H., Perkins, K. A., Clarke, E. B., Dunn, B., Gerber, A. H., & Morrow, E. M. (2020). Autism heterogeneity in a densely sampled U.S. population: Results from the first 1,000 participants in the RI-CART study. *Autism Research*, 13(3), 474–488. <https://doi.org/10.1002/aur.2261>

McDaniel, J. (2017). *Child reciprocal vocal contingency measure using automated vocal analysis with children with autism spectrum disorder*. International Society for Autism Research.

McDaniel, J., Yoder, P., Crandall, M., Millan, M. E., Ardel, C. M., Gengoux, G. W., & Hardan, A. Y. (2020). Effects of pivotal response treatment on reciprocal vocal contingency in a randomized controlled trial of children with autism spectrum disorder. *Autism*, 24(6), 1566–1571. <https://doi.org/10.1177/1362361320903138>

McDaniel, J., Yoder, P., Estes, A., & Rogers, S. J. (2020). Predicting Expressive Language From Early Vocalizations in Young Children With Autism Spectrum Disorder: Which Vocal Measure Is Best? *Journal of Speech, Language, and Hearing Research*, 63(5), 1509–1520. [https://doi.org/10.1044/2020\\_JSLHR-19-00281](https://doi.org/10.1044/2020_JSLHR-19-00281)

Meera, S. S. (2019). *Canonical Babbling in 9-month-old infants later diagnosed with autism spectrum disorder: A naturalistic evaluation of all-day recordings*. International Society for Autism Research.

Messinger, D., Young, G. S., Ozonoff, S., Dobkins, K., Carter, A., Zwaigenbaum, L., Landa, R. J., Charman, T., Stone, W. L., Constantino, J. N., Hutman, T., Carver, L. J., Bryson, S., Iverson, J. M., Strauss, M. S., Rogers, S. J., & Sigman, M. (2013). Beyond autism: A baby siblings research consortium study of high-risk children at three years of age. *Journal of the American Academy of Child & Adolescent Psychiatry*, 52(3), 300–308.e1. <https://doi.org/10.1016/j.jaac.2012.12.011>

Millager, R. A., Feldman, J., Williams, Z., Shibata, K., Martinez-Torres, K. A., Bryan, K. M., Pruitt, D. G., Mitchell, J. T., Markfeld, J., Merritt, B., Daniels, D. E., Jones, R. M., & Woynaroski, T. G. (2024). Diversity of research participant gender, race, and ethnicity in communication sciences and disorders: A systematic review of ASHA publications in 2020. *Perspectives of the ASHA Special Interest Groups*, 9(3), 836–852. [https://doi.org/10.1044/2024\\_PERSP-23-00204](https://doi.org/10.1044/2024_PERSP-23-00204)

Moffitt, J. M., Ahn, Y. A., Custode, S., Tao, Y., Mathew, E., Parlade, M., Hale, M., Durocher, J., Alessandri, M., Perry, L. K., & Messinger, D. S. (2022). Objective measurement of vocalizations in the assessment of autism spectrum disorder symptoms in preschool age children. *Autism Research*, 15(9), 1665–1674. <https://doi.org/10.1002/aur.2731>

Oller, D. K. (2015, May 14). *Towards improved clinically useful automated vocal assessments for the prediction of ASD*. International Society for Autism Research.

Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., Yapaner, U., & Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences*, 107(30), 13354–13359. <https://doi.org/10.1073/pnas.1003882107>

Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, 10(3), 233–240. <https://doi.org/10.1111/jebm.12266>

Paruthi, S., Brooks, L. J., D'Ambrosio, C., Hall, W. A., Kotagal, S., Lloyd, R. M., Malow, B. A., Maski, K., Nichols, C., Quan, S. F., Rosen, C. L., Troester, M. M., & Wise, M. S. (2016). Recommended amount of sleep for pediatric populations: A consensus statement of the American academy of sleep medicine. *Journal of Clinical Sleep Medicine: JCSM: Official Publication of the American Academy of Sleep Medicine*, 12(6), 785–786. <https://doi.org/10.5664/jcsm.5866>

Pawar, R., Albin, A., Gupta, U., Rao, H., Carberry, C., Hamo, A., Jones, R. M., Lord, C., & Clements, M. A. (2017). Automatic analysis of LENA recordings for language assessment in children aged five to fourteen years with application to individuals with autism. In *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)* (pp. 245–248). <https://doi.org/10.1109/BHI.2017.7897251>

Pellicano, E., Dinsmore, A., & Charman, T. (2014). What should autism research focus upon? Community views and priorities from the United Kingdom. *Autism*, 18(7), 756–770. <https://doi.org/10.1177/1362361314529627>

Rankine, J., Li, E., Lurie, S., Rieger, H., Fourie, E., Siper, P. M., Wang, A. T., Buxbaum, J. D., & Kolevzon, A. (2017). Language ENvironment Analysis (LENA) in Phelan-McDermid Syndrome: Validity and suggestions for use in minimally verbal children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 47(6), 1605–1617. <https://doi.org/10.1007/s10803-017-3082-8>

Räsänen, O., Seshadri, S., Lavechin, M., Cristia, A., & Casillas, M. (2021). ALICE: An open-source tool for automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings. *Behavior Research Methods*, 53(2), 818–835. <https://doi.org/10.3758/s13428-020-01460-x>

Richards, J. A., Gilkerson, J., Paul, T. D., & Xu, D. (2008). *The LENA Automatic Vocalization Assessment* (Technical Report LTR-08-1; LENA Technical Report). LENA Foundation.

Richards, J. A., Xu, D., & Gilkerson, J. (2010). *Development and performance of the LENA automated autism screen* (Technical Report LTR-10-1; LENA Technical Report). LENA Foundation.

Sabatos-DeVito, M. G. (2016). *Relations between computerized LENA recordings of conversational turns and lab-based measures of social engagement in children with autism spectrum disorder*. International Society for Autism Research.

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3E [3rd ed.]). SAGE.

Sandbank, M., Bottema-Beutel, K., Crowley, S., Cassidy, M., Dunham, K., Feldman, J. I., Crank, J., Albaran, S. A., Raj, S., Mahbub, P., & Woynaroski, T. G. (2020). Project AIM: Autism intervention meta-analysis for studies of young children. *Psychological Bulletin*, 146(1), 1–29. <https://doi.org/10.1037/bul00000215>

Seidl, A., Cristia, A., Soderstrom, M., Ko, E.-S., Abel, E. A., Kellerman, A., & Schwichtenberg, A. J. (2018). Infant–mother acoustic–prosodic alignment and developmental risk. *Journal of Speech, Language, and Hearing Research*, 61(6), 1369–1380. [https://doi.org/10.1044/2018\\_JSLHR-S-17-0287](https://doi.org/10.1044/2018_JSLHR-S-17-0287)

Shattuck, P. T., Durkin, M., Maenner, M., Newschaffer, C., Mandell, D. S., Wiggins, L., Lee, L.-C., Rice, C., Giarelli, E., Kirby, R., Baio, J., Pinto-Martin, J., & Cuniff, C. (2009). Timing of identification among children with an autism spectrum disorder: Findings from a population-based surveillance study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 48(5), 474–483. <https://doi.org/10.1097/CHI.0b013e31819b3848>

Spencer, T. D., Tolentino, T. J., & Foster, M. E. (2023). Impact of discourse type and elicitation task on language sampling outcomes. *American Journal of Speech-Language Pathology*, 32(6), 2827–2845. [https://doi.org/10.1044/2023\\_AJSLP-22-00365](https://doi.org/10.1044/2023_AJSLP-22-00365)

Steinbrenner, J. R., McIntyre, N., Rentschler, L. F., Pearson, J. N., Luelmo, P., Jaramillo, M. E., Boyd, B. A., Wong, C., Nowell, S. W., Odom, S. L., & Hume, K. A. (2022). Patterns in reporting and participant inclusion related to race and ethnicity in autism intervention literature: Data from a large-scale systematic review of evidence-based practices. *Autism*. <https://doi.org/10.1177/13623613211072593>

Stolte, M. (2017). *Bridging the research to practice gap in autism spectrum disorder treatment*. University of Alberta.

Sulek, R., Smith, J., Bent, C. A., Hudry, K., Trembath, D., Vivanti, G., & Dissanayake, C. (2022). The utility of LENA as an indicator of developmental outcomes for young children with autism. *International Journal of Language & Communication Disorders*, 57(1), 103–111. <https://doi.org/10.1111/1460-6984.12678>

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., & for the 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>

Swanson, M. R., Shen, M. D., Wolff, J. J., Boyd, B., Clements, M., Rehg, J., Elison, J. T., Paterson, S., Parish-Morris, J., Chappell, J. C., Hazlett, H. C., Emerson, R. W., Botteron, K., Pandey, J., Schultz, R. T., Dager, S. R., Zwaigenbaum, L., Estes, A. M., Piven, J., & Gu, H. (2018). Naturalistic language recordings reveal “hypervocal” infants at high familial risk for autism. *Child Development*, 89(2). <https://doi.org/10.1111/cdev.12777>

Tager-Flusberg, H. (2006). Defining language phenotypes in autism. *Clinical Neuroscience Research*, 6(3–4), 219–224. <https://doi.org/10.1016/j.cnr.2006.06.007>

Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum: Minimally verbal children with ASD. *Autism Research*, 6(6), 468–478. <https://doi.org/10.1002/aur.1329>

Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., Rice, M., Stoel-Gammon, C., Wetherby, A., & Yoder, P. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with Autism Spectrum Disorders. *Journal of Speech, Language, and Hearing Research*, 52(3), 643–652. [https://doi.org/10.1044/1092-4388\(2009/08-0136\)](https://doi.org/10.1044/1092-4388(2009/08-0136)

Trembath, D., Westerveld, M. F., Teppala, S., Thirumanickam, A., Sulek, R., Rose, V., Tucker, M., Paynter, J., Hetzroni, O., Keen, D., & Vivanti, G. (2019). Profiles of vocalization change in children with autism receiving early intervention. *Autism Research*, 12(5), 830–842. <https://doi.org/10.1002/aur.2075>

Vivanti, G., Dissanayake, C., Duncan, E., Feary, J., Capes, K., Upson, S., Bent, C. A., Rogers, S. J., Hudry, K., Jones, C., Bajwa, H., Marshall, A., Maya, J., Pye, K., Reynolds, J., Rodset, D., & Toscano, G. (2019). Outcomes of children receiving Group-Early Start Denver Model in an inclusive versus autism-specific setting: A pilot randomized controlled trial. *Autism*, 23(5), 1165–1175. <https://doi.org/10.1177/1362361318801341>

Wang, Y., Hartman, M., Aziz, N. A., Arora, S., Shi, L., & Tunison, E. (2017). A systematic review of the use of LENA technology. *American Annals of the Deaf*, 162(3), 295–311. <https://doi.org/10.1353/aad.2017.0028>

Wang, Y., Williams, R., Dilley, L., & Houston, D. M. (2020). A meta-analysis of the predictability of LENATM automated measures for child language development. *Developmental Review: DR*, 57, 100921. <https://doi.org/10.1016/j.dr.2020.100921>

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, 25(7), 1314–1324. <https://doi.org/10.1177/0956797614531023>

Warren, S. F., Gilkerson, J., Richards, J. A., Oller, D. K., Xu, D., Yapanel, U., & Gray, S. (2010). What automated vocal

analysis reveals about the vocal production and language learning environment of young children with autism. *Journal of Autism and Developmental Disorders*, 40(5), 555–569. <https://doi.org/10.1007/s10803-009-0902-5>

Watson, L. R., & Crais, E. R. (2013). Translating between research and practice in serving infants at risk for ASD. *Perspectives on Language Learning and Education*, 20(1), 4–13. <https://doi.org/10.1044/lle20.1.4>

Winters, K. L., Jasso, J., Pustejovsky, J. E., & Byrd, C. T. (2022). Investigating narrative performance in children with developmental language disorder: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, 65(10), 3908–3929. [https://doi.org/10.1044/2022\\_JSLHR-22-00017](https://doi.org/10.1044/2022_JSLHR-22-00017)

Wittke, K., Masturgeorge, A. M., Ozonoff, S., Rogers, S. J., & Naigles, L. R. (2017). Grammatical language impairment in autism spectrum disorder: Exploring language phenotypes beyond standardized testing. *Frontiers in Psychology*, 8. <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00532>

Woynaroski, T., Oller, D. K., Keceli-Kaysili, B., Xu, D., Richards, J. A., Gilkerson, J., Gray, S., & Yoder, P. (2017). The stability and validity of automated vocal analysis in preverbal preschoolers with autism spectrum disorder: Automated vocal analysis in ASD. *Autism Research*, 10(3), 508–519. <https://doi.org/10.1002/aur.1667>

Xu, D., Richards, J. A., & Gilkerson, J. (2014). Automated analysis of child phonetic production using naturalistic recordings. *Journal of Speech, Language, and Hearing Research*, 57(5), 1638–1650. [https://doi.org/10.1044/2014\\_JSLHR-S-13-0037](https://doi.org/10.1044/2014_JSLHR-S-13-0037)

Yoder, P. J., Lloyd, B. P., & Symons, F. J. (2018). *Observational measurement of behavior* (2nd ed.). Brookes Publishing.

Yoder, P. J., Oller, D. K., Richards, J. A., Gray, S., & Gilkerson, J. (2013). Stability and validity of an automated measure of vocal development from day-long samples in children with and without autism spectrum disorder: Stability and validity of vocal measure. *Autism Research*, 6(2), 103–107. <https://doi.org/10.1002/aur.1271>

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scales* (5th ed.). The Psychological Corporation. <https://doi.org/10.1037/t15141-0003994>

Zuckerman, K., Lindly, O. J., & Chavez, A. E. (2017). Timeliness of autism spectrum disorder diagnosis and use of services among U.S. elementary school-aged children. *Psychiatric Services (Washington, D.C.)*, 68(1), 33–40. <https://doi.org/10.1176/appi.ps.201500549>

Zwaigenbaum, L., Bryson, S. E., Brian, J., Smith, I. M., Roberts, W., Szatmari, P., Roncadin, C., Garon, N., & Vaillancourt, T. (2016). Stability of diagnostic assessment for autism spectrum disorder between 18 and 36 months in a high-risk cohort. *Autism Research*, 9(7), 790–800. <https://doi.org/10.1002/aur.1585>