Journal of Child Language (2025), 1–24
doi:10.1017/S0305000924000606

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

An Evaluation of LENA Start™ Using Measures Derived from Parent–Child Interactions

Kirstin Kuchler[1], Marianne Elmquist[2], Scott R. McConnell[3] and Lizbeth H. Finestack[1]

[1]Speech-Language-Hearing Sciences, University of Minnesota-Twin Cities; [2]Waisman Center, University of Wisconsin-Madison and [3]Educational Psychology, University of Minnesota-Twin Cities
Corresponding author: Lizbeth H. Finestack; Email: finestac@umn.edu

(Received 08 December 2022; revised 14 June 2024; accepted 10 September 2024)

Abstract

It is important to assess learning in both familiar and unfamiliar conditions to determine the extent of learning generalisation. In this study, we evaluated parent language outcomes of LENA Start™, a parent-implemented intervention, using distal measures derived from a parent–child free play interaction.

Forty-four parents and their child (mean: 20.8 months) participated in LENA Start™ or early childhood education curriculum intervention, in either English (n = 34) or Spanish (n = 10). We completed between- and within-group analyses using language and communication behaviour measures derived from parent–child interactions elicited outside the home with researcher-supplied materials (i.e., distal measures).

Group comparisons revealed significant differences on a subset of behavioural measures. Spanish-speaking parents in the LENA Start™ group demonstrated significant gains on three measures. While LENA Start™ has been associated with parent gains on proximal language measures, the results of this study reveal limited generalisation of skills based on distal measures.

Abstract in spanish

Es importante evaluar el aprendizaje tanto en condiciones familiares como en condiciones no familiares. En el presente estudio se evaluaron los resultados relacionados con el lenguaje paterno de LENA Start, una intervención implementada por padres/madres, usando medidas distales obtenidas de interacciones de juego no estructurado entre padre e hijo. Cuarenta y cuatro padres y sus hijos (promedio: 20.8 meses) participaron en LENA Start o una intervención del currículo de Educación Infantil en inglés (n = 34) o en español (n = 10). Se llevaron a cabo varios análisis entre grupos y dentro de los grupos utilizando medidas de comportamiento lingüístico y comunicativo derivadas de interacciones entre padre e hijo realizadas fuera de la casa con materiales provistos por las investigadoras (es decir, medidas distales). Las comparaciones de grupo revelaron diferencias significativas en un subconjunto de las medidas de comportamiento. Los padres hispanohablantes en el grupo de LENA Start

ganaron en tres de las medidas. Mientras LENA Start ha sido asociado con ganancias paternas en medidas lingüísticas proximales, los resultados del presente estudio revelan generalización de habilidades limitada basada en medidas más distales.

Introduction

The language environment and experiences in a child's first 3 years of life greatly impact the child's overall language development (Hirsh-Pasek et al., 2015; Hoff, 2006). Parents and caregivers are often the first source of language input for a child. Given that parents do not typically receive instruction on language development, many parents may benefit from additional and targeted support to build the capacity of knowledge of child language development and skills to enhance their child's language development. Improving parent knowledge in these areas is critical to help minimise potential language delay or even language disorders. A variety of language development strategies (e.g., dialogic reading, modelling, joint attention) have been evaluated to support parents in improving their child's language environment (see Roberts et al., 2019; Finestack et al., 2022). These interventions have various levels of evidence supporting their efficacy.

In this study, we examined outcomes associated with LENA Start™, a community-based intervention aimed to support children's language environment by increasing their parents' knowledge and implementation of language-facilitating strategies. Previous evaluations (Beecher & Van Pay, 2019; Beecher & Van Pay, 2020; Elmquist et al., 2021), including our own, have primarily evaluated the effectiveness of LENA Start™ using proximal measures directly associated with the intervention. As a follow-up of Elmquist et al.' (2021) study, in this study, we evaluated LENA Start™ using more distal adult language measures associated with a simulated parent–child play session.

Environmental Factors Affect Language Development

Two frameworks can help structure the discussion of caregiver–child communication: the bioecological framework (Bronfenbrenner & Ceci, 1994) and the transactional model framework (Sameroff, 2009). The bioecological framework centers the caregiver–child communication environment within factors and experiences that affect and are affected by caregiver–child relationships. Thus, the bioecological framework is useful to help understand how various environmental factors (e.g., cultural, socio-economic status) contribute to a child's language development (Hoff, 2006).

The bioecological framework provides evidence for the functional relationship between a specific behaviour and the environmental factors that affect an individual's behaviour (Ford et al., 2020). The relationship is affected by the proximity of the specific behaviour to the environmental factors. More specifically, the caregiver–child communication environment is closely affected by the caregiver's communication beliefs, knowledge, and behaviour and distantly affected by policies and practices of the larger community (Ford et al., 2020). By saying that language development is highly impacted by a child's environment, we assume that language learning occurs in response to each of these factors to some degree.

The second framework is the transactional model. Transactional communication is dynamic and changes based on communication participants' reactions within an interaction (Sameroff, 2009). A transactional experience requires each entity to adapt to the other. Changes within the child's environment affect developmental changes and

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

developmental changes affect the child's environment "as a function of their mutual influence on one another." (Sameroff, 2009, p. 7). Caregivers control the child's environment and can therefore direct the changes.

These frameworks situate the quality of a parent's interaction with their child as an important element of development that changes over time. Parents serve the role of providing opportunities and reinforcements for their child's language (Ford et al., 2020). Thus, instructional and intervention programmes exist to help teach parents strategies to support their child's communication and language development. Moreover, there is a growing body of research investigating the efficacy and effectiveness of such programmes. For example, in a recent scoping review, Finestack et al. (2022) identified 59 studies that evaluated the efficacy of caregiver-implemented interventions for children 0 to 48 months of age identified as having language impairment. Similarly, a meta-analysis conducted by Heidlage et al. (2019) included 25 randomised controlled trials of parent-implemented language interventions for children 0 through 8 months of age with language impairment, at-risk for language impairment, or who were developing typically. The findings of Heidlage et al. indicated significant positive treatment effects on parent use of language-facilitating behaviours; however, only five studies measured parent language; the others exclusively measured child language use. Thus, more research is needed to understand the impact of parent education and intervention to support their child's communication environment. One parent-implemented intervention programme for which a body of evidence is beginning to mount (i.e., Beecher & Van Pay, 2019; Beecher & Van Pay, 2020; Elmquist et al., 2021) is the LENA Start™ Programme.

LENA Start™ Programme

LENA (Language Environment Analysis), a national non-profit company, designed programmes to promote child language development by improving parent knowledge and implementation of language-facilitating strategies. These programmes include parent education (i.e., LENA Start™), childcare professional development (i.e., LENA Grow™), and a home-visiting language development programme (i.e., LENA Home™). Prior to creating these programmes, the LENA company developed technology that records and synthesises features of child vocalisations and adult–child conversations (i.e., LENA® Digital Language Processor recording device). This technology is a key component of the LENA education programmes.

LENA Start™ is a 13-week programme that teaches parents about child language development and facilitation strategies (see Appendix A for weekly summary of curriculum). Parents participating in the LENA Start™ programme receive a LENA® Digital Language Processor recording device. This device is worn by their child once each week across the 13-week intervention period to record 16 hours of continuous audio. Four measures that are derived from the day-long LENA recordings, using proprietary algorithms, include the number of adult words heard by the child (AWC), the number of vocalisations that the child produced (CVC), the number of child–adult conversational turns (CTC), and minutes of electronic noise. Interventionists review each measure with parents on a weekly basis and encourage parents to use targeted facilitation strategies (i.e., Talking Tips; LENA Foundation, 2015) to increase their daily rates of AWC and support CVCs and CTCs.

There are four known recent studies in which researchers have evaluated the LENA Start™ programme (i.e., Beecher & Van Pay, 2019, which includes two studies; Beecher & Van Pay, 2020; Elmquist et al., 2021). In each of these studies, researchers evaluated the

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

intervention using LENA measures that the interventionists reviewed with parents weekly. For example, to increase the AWC, the interventionist would encourage parents to increase the number of words directed to their child by getting down at their child's level, commenting on what their child is doing, and naming the child's interests (LENA Foundation, 2015). To increase CTC, the interventionist might suggest to the parent to wait for the child responses and then take turns.

Beecher and Van Pay (2019) conducted two studies evaluating the effects of LENA Start™: a pilot study and a follow-up study after making programmatic changes. All participants in the studies received the intervention and researchers analysed pre- and post-intervention performance. Researchers also examined the impact of socio-economic status on outcomes. In the first study, which evaluated an 8-week LENA Start™ programme, parents significantly increased their AWC by 4,000 words a day. This increase was associated with a large effect size (d = 0.99). In this study, Beecher and Van Pay also measured change in parent knowledge of language development using the Survey of Parent Expectations and Knowledge About Language Learning (SPEAK; Suskind et al., 2016). Pre-intervention SPEAK measures revealed a significant difference in parent knowledge of child language development between parents who received WIC (a federal assistance programme for low-income pregnant and/or breastfeeding women and children under 5 years old) and those who did not. However, these within-group differences were not significant post-intervention, suggesting an effect of intervention for parents who received WIC.

In their second study, Beecher and Van Pay (2019) evaluated an expanded 13-week LENA Start™ programme (also evaluated by Elmquist et al., 2021) with a new cohort of participants. They found significant increases, between pre- and post-intervention, for both AWC (increase of 912 words) and CTC (increase of 37 conversational turns). There was also a significant increase pre- to post-intervention in parent knowledge of child language development, measured by SPEAK, which was associated with a large effect size (d = 0.94).

In a third study, Beecher and Van Pay (2020) collected data from a comparison group and compared their results to the group that received treatment in the Beecher and Van Pay (2019) Study 2. Intervention in the 2019 study was conducted in three libraries; the comparison group in the 2020 study was recruited from two of these same libraries. The 28 comparison families were matched with 28 intervention families for analysis. The comparison group met researchers in the library weekly for 6 weeks to obtain a LENA recording device that their child wore for a full day but received no other instruction. Results revealed that the intervention group demonstrated greater improvement than the comparison group over the course of intervention based on the LENA measures of AWC, CTC, and CVC. The AWC comparison was associated with a small effect size (d = 0.36); the CTC effect size was large (d = 0.80); and the CVC effect size was medium (d = 0.67).

Elmquist et al. (2021) also compared performance on LENA measures between a group receiving the LENA Start™ intervention and a comparison group participating in a typical parent education Early Childhood Family Education (ECFE) class. Results revealed significant group differences in CTC with a large effect size (g = 0.81). There was not a significant difference in CVC (no effect size reported) between treatment and comparison groups. Elmquist et al. also examined parents perception of their child's language using The Developmental Snapshot (Snapshot; LENA EarlyTalk, 2022) once at pre-intervention and at post-intervention. The Snapshot is a tool created by LENA Start™ to evaluate parents' perception of their child's progress. Results revealed a significant, medium effect size (g = 0.57) when the LENA Start™ and comparison groups were compared. However, when controlling for child age and family income below the poverty line, results no longer differed significantly.

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Beecher and Van Pay (2019 and 2020) and Elmquist et al. (2021) both found that significant pre-treatment to post-treatment differences and large effect size changes for CTC. Beecher and Van Pay (2020) found a significant difference and medium effect size for CVC, which Elmquist et al. (2021) did not find. To date, however, evaluation of LENA Start™ effects has been primarily restricted to measures directly targeted in the intervention (i.e., LENA AWC, CTC, CVC measures) or broad parent report measures (i.e., SPEAK, Snapshot), all of which are measures, that interventionists specifically taught caregivers how to modify and knowledge of strategies explicitly targeted in the intervention. A further question is whether the LENA Start™ programme effects generalise to measures less closely tied to the intervention itself.

Distal and proximal measures

Intervention outcomes can measure a variety of skills, including skills that are directly targeted in the intervention and those that are not specifically targeted. Additionally, these skills can be measured in contexts that closely reflect the context in which the intervention took place or contexts that vary from the intervention context. Measures based on contexts that are less like the intervention context are more likely to evaluate the generalisability of intervention targets. Given that the goal of parent-implemented interventions is for the parent to learn to use language-facilitation strategies in a variety of environments outside of the direct intervention environment, it is important to evaluate learning with a variety of measures.

One way to categorise intervention outcome measures is based on their proximity and boundness to the intervention target (Yoder et al., 2013; Sandbank et al., 2021). Proximity refers to the overlap of an outcome with intervention target(s) (e.g., measuring outcomes not directly taught in the intervention). Boundness refers to the extent that an outcome reflects change outside of the immediate intervention context (e.g., different materials, setting, partner). Although measures should be considered on a continuum, these categories can yield four distinct types of contexts. Proximal, context-bound measures are those that are taught and measured within the intervention context. Proximal, generalised measures are those that are directly taught by the intervention and measured in a context outside of direct intervention (e.g., different settings or materials). Distal, context-bound measures are related to the intervention target but not directly taught and measured in the same context as the intervention (e.g., measuring initiating joint attention when the intervention taught responding to joint attention, using the same setting, materials, communication partner as the intervention). Distal, generalised measures are not directly taught in the intervention and are measured in a context outside of the intervention (e.g., initiating joint attention in a free play session with peers after a clinician-led intervention session that taught responding to joint attention). Yoder et al. (2013) identified a 63% likelihood of finding a treatment effect with proximal measures and only a 39% likelihood with distal measures. Additionally, context-bound outcomes were associated with an 82% likelihood of finding a treatment effect compared to 33% with generalised outcomes. Yoder et al. argued that while clinicians may not be surprised by these results due to their anecdotal experiences, investigators need to carefully select and define intervention outcome measures. The most robust interventions would be those that indicate positive outcomes on both proximal and distal measures.

Beecher and Van Pay (2019 & 2020) and Elmquist et al. (2021) primarily used proximal measures to evaluate the LENA Start™ programme outcomes. These studies

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

evaluated adult language use outcome measures that were collected using the LENA recorder (i.e., AWC, CVC, and CTC). Parents learned about these measures during the structured curriculum and received direct feedback on their use. The LENA measures are considered proximal and their use to evaluate the intervention may be considered a "teach to the test" approach. These studies did not evaluate adult language outcomes using more distal measures not directly associated with the LENA Start™ curriculum. The distinction between proximal and distal effects is closely related to issues of treatment generalisation (e.g., Stokes & Baer, 1977). Treatment effects can be evaluated both directly and across nontreated contextual variables (e.g., time, setting, behaviour, persons). Like distal effects (c.f., Yoder et al., 2013), generalisation effects cannot be assumed when treatment effects are obtained; however, generalisation effects may often better represent the broad or clinical intent of intervention.

To more rigorously evaluate LENA Start™, it is necessary to examine secondary outcome measures that are more distal to the intervention as the LENA measures (i.e., AWC, CVC, and CTC) are the closest possible measures in proximity to the LENA Start™ curriculum. Distal measures can indicate if the intervention generalises to other behaviours, settings, or stimuli. Thus, this study serves as a follow-up to the Elmquist et al. (2021) study by examining intervention outcomes based on more distal measures derived from a 6-min parent–child free play interaction conducted outside of families' homes. Further examination of LENA Start™ using more distal measures will allow us to better understand the impact of the intervention.

Current Study Research Questions

The purpose of this study was to determine if parents implemented the strategies they learned through LENA Start™ in a simulated naturalistic environment. We compared the outcomes of parents who participated in a LENA Start™ intervention group to parents enrolled in a general early childhood education group that comprised a comparison group. As a follow-up to the Elmquist et al. (2021) study, we examined more distal parent communication outcome measures derived from 6-min parent–child play interactions completed outside the home with toys provided before and after intervention. We derived two types of measures from the play interactions: those based on transcripts of videotaped interaction and analysed using measures derived from the Systematic Analysis of Language Transcripts (SALT; Miller & Iglesias, 2016) and those based on behavioural coding of these same interactions. The measures derived from SALT included the number of total words (NTW), number of different words (NDW), mean length of utterance by morpheme for English speakers (MLU-M), mean length of utterance by words for Spanish speakers (MLU-W), number of total utterances (NumUtt), and mean turn length of utterance (MTLUtt). The measures derived from behavioural coding included maintaining the child's communication topic (maintain), redirecting the child's attention (redirect), commenting to the child (comment), requesting the child to say something (request for verbal reply), introducing a new topic (introduce), reading to the child (reading), and requesting the child to do something (request for behavioural comply).

Previous studies (i.e., Beecher & Van Pay, 2019; Beecher & Van Pay, 2020; Elmquist et al., 2021) analyses are considered proximal, generalised measures (Yoder et al., 2013). The interventionist teaches parents, in a community location within a group setting to speak more with their child and then evaluates the results in the daily routines of the parent and child (e.g., home and/or out and about in the community) using three main

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

gross measures (i.e., AWC, CVC, and CTC) which simply reveal if more language is used after intervention than before. Most of the measures in this study are considered distal, generalised because they evaluate language use at a more fine-grained level outside of the intervention context. For example, NDW is a more nuanced measure of language use compared to AWC. Additionally, in this study, we measured language behaviour not directly taught in the LENA Start™ curriculum (e.g., MLU or sentence complexity). We were specifically interested in evaluating measures that reflect the quality of parent–child interactions in addition to quantity.

Like Elmquist (2021), we analysed three different demographic groups: the Full Sample, English-only speaking, and Spanish-only speaking. We organised the sets of results in this manner because, although all participants were involved in the same cohort of data collection, English and Spanish versions of LENA Start™ have different language expectations and structures and therefore should also be considered separately. The specific study questions were as follows:

Research Question 1: Are there significant group differences in adult language change between the Full Sample (English- and Spanish-speaking) of the LENA Start™ programme participants and the Comparison group participants based on SALT measures?

Research Question 2: Are there significant within-group pre- to post-intervention differences in adult language for the English- or Spanish-speaking participants in the LENA Start™ programme based on SALT measures?

Research Question 3: Are there significant group differences in change in adult communication between the Full Sample (English- and Spanish-speaking) participants in the LENA Start™ programme and participants in the Comparison group based on coded communicative behaviour measures?

Research Question 4: Are there significant within-group pre- to post-intervention differences in adult communication for the English- and Spanish-speaking participants in the LENA Start™ programme based on coded communicative behaviour measures?

Although more distal to the intervention, we predicted that both the English- and Spanish-speaking parents who received the LENA Start™ intervention would demonstrate significant differences on both the SALT measures and communicative behaviour measures compared to the Comparison group as well as significant pre- to post-intervention differences because the programme teaches parents to increase language use with their children. For the communicative behaviour measures, we predicted significant increases in behaviours that promote language development, including maintain, comment, and verbal reply. We expected behaviours that impede the child's opportunity to communicate back to the parent to decrease, including redirect, introduce, reading, and request for behavioural comply. The Reading communication behaviour code in this study does not refer to shared book reading which is very predictive and supportive of language development (Noble et al., 2020). The code refers to the parent reading the exact words from the book and not new language produced by either the parent or child about the book or reading experience, which would be done using a dialogic reading approach.

Method

Recruitment and participants

After receiving institutional review board approval, we recruited participants across four partnering Early Childhood Family Education (ECFE) Programs in the Midwest of the United States. ECFE is a publicly funded programme, typically operated by local

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

education agencies, offering a range of interaction and educational opportunities to parents and young children. Classes are taught by a licensed parent educator, typically working with a licensed early educator and/or educational assistants. The study required that child participants were between the ages of 0 and 2 years, 11 months. Parents chose to participate in an ECFE programme delivering either LENA Start™ curriculum (four classrooms in English and one classroom in Spanish) or regular parent education classes (two classrooms in English and one classroom in Spanish). Each ECFE site agreed to accept community-wide enrolment in response to recruitment for this project.

Recruitment of participants included community-wide marketing of typically offered ECFE classes in addition to LENA Start™ classes. Additionally, some parents enrolled in other ECFE classes or district programmes were invited to participate in the research. Participants' enrolment in scheduled ECFE classes pre-determined assignment to the LENA Start™ group or the Comparison group. We cannot rule out that parents may have intentionally signed up for LENA Start™ classes over regular ECFE classes. Therefore, the group assignment was not random.

Participants were recruited in the fall of 2017. Funding lasted only for recruitment and participation for one semester, which is one reason for the small sample size. A total of 62 families (LENA Start™: $n = 41$; Comparison: $n = 21$) consented to participate and complete the protocol found in Elmquist et al. (2021); 44 of those participants consented to participate in this study. The current study included 34 English-speaking dyads (LENA Start™ $n = 19$, Comparison $n = 15$) and 10 Spanish-speaking dyads (LENA Start™ $n = 9$, Comparison $n = 1$) (see Table 1) who participated in both pre- and post-intervention parent–child free play interactions. All families (LENA Start™ and Comparison) were compensated with a $30 Target gift card for their time when they returned the materials and completed the assessments. Families in the LENA Start™ group also received an age-appropriate reading book each week throughout the programme. At the time of the study, the required materials and resources for implementation of the LENA Start™ programme cost approximately $270 per family.

Intervention procedures

The LENA Start™ and Comparison parent–child dyads participated in a 13-week parent education programme that included 30 minutes of parent–child activity time and then separate child and parent group classes. When separated, the children participated in early childhood education classes while parents completed ECFE classes in which the LENA Start™ group focused on language development as described by the program's standard intervention manual, and the Comparison group focused on general child development. For both groups, licensed ECFE teachers conducted the parent education classes. The LENA Start™ teachers received specialised training to conduct the language-focused intervention curriculum.

In addition to completing weekly LENA recordings, parents participating in LENA Start™ learned about the importance of language development. The educators provided parents with 14 Talking Tips, which included easy-to-remember strategies that encourage frequent, quality interactions between parents and their children. During the classes, parents had the opportunity to create goals, share tips with one another on incorporating strategies, and build social capital further promoting the use of language skills. During these group sessions, the educator gave parents a LENA report which visually displayed their last LENA recording, specifically providing feedback on AWC, CVC, and CTC. If parents

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

wanted to discuss their reports with parent educators they could do so, but it was not embedded into the programme. The primary goal of the LENA Start™ programme is for parents and their children to make weekly progress on AWC, CVC, and CTC measures.

When Comparison group parents separated from their children, they attended regular ECFE programming, which included broad parenting information on topics such as toilet training, sleep, behaviour, and language development. Participants in the Comparison group completed two LENA recordings: once at the beginning of their 13-week traditional ECFE programme and once at the end. They received the LENA recording reports for both recordings at the conclusion of the study.

LENA technology. The LENA® Digital Language Processor is a small digital recording device created to record a child's language environment (Richards et al., 2017). The LENA Start™ and Comparison group dyads used the Digital Language Processor for up to 16 hours during a single day to record the child's language environment. Recordings occurred weekly for participants in the LENA Start™ group and pre- and post-intervention for participants in the Comparison group. On each recording day, the child wore a specialised vest in which the device was placed. The educators instructed parents to leave the recorder on all day. Parent and child language was analysed using automated computer algorithms (Greenwood et al., 2010) to measure AWC, CVC, CTC, and the number of minutes of electronic noise. The results of the LENA-derived pre- and post-intervention measures were presented in Elmquist et al. (2021).

Parent–child Interaction

Parents in both the LENA Start™ and Comparison groups engaged in a free play interaction with their child before and after the intervention period. Video recordings of the interaction were conducted at the participating ECFE programmes in a convenient location (typically an empty classroom or observation room). Researchers provided parents with a standardised set of age-appropriate toys and books. Toys were the same for both Spanish- and English-speaking families and included a barn set, doll, telephone, stuffed puppy, light-up rattle, and shape sorter. For English-speaking dyads, the two books included "The Very Hungry Caterpillar" (Carle, 1969) and "Touch and Feel: Baby Animals" (Dorling Kindersley, Inc., 2008). For the Spanish-speaking dyads, the two books included "La Oruga Muy Hambrienta" ("The Very Hungry Caterpillar") (Carle, 2002) and "Buenos Días, Bebé!" ("Good Morning, Baby!") (DK, 2004). Before researchers started recording parents interacting with their child, a researcher provided an overview of the procedures to each dyad (that they would be observed interacting as they typically would at home with the toys and books provided for 7–10 minutes). This protocol is a modified version on the protocol used by the British Autism Study of Infant Siblings (BASIS; Wan et al., 2012) For the Spanish-speaking families, this overview was translated to Spanish by ECFE staff. No warm-up play opportunity was given as the goal of distal measures is to understand how the intervention has been generalised into other contexts. During the video recording, the researcher refrained from interacting with the dyads. Each interaction lasted between 7 and 10 minutes.

Transcription procedures, reliability, and measures

To account for different lengths in recording, the first 6 minutes of each observation was transcribed and coded for analysis. This is in line with other parent–child interaction

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

research. For example, Wan and colleagues (2012) researched interactions among parents and children 6–10 months old using a 6-minute interaction and found similar results of child's level of activity as self-report. In addition, the early communication indicator (ECI) is a measure of a child's communicative behaviour based on a 6-minute play-based observation with a familiar communication partner (Luze et al., 2001). The videos of the English-speaking dyads were transcribed independently by undergraduate and PhD level students in an academic speech-language-hearing sciences department using the standard English conventions for the Systematic Analysis of Language Transcripts (SALT; Miller & Iglesias, 2016). Transcriptionists were blind to study details, group assignments, and knowledge of whether the interaction was recorded pre- or post-intervention. The training procedure included a review of the SALT transcription manual, completing SALT transcription tutorials available through the SALT website, completing the training transcripts, and weekly lab meetings during which transcriptionists discussed questions. Interrater reliability between the individual transcriptions and the master transcriptions was required to be 80% or above prior to the transcriptionist independently transcribing videos used in data analysis. We calculated reliability using percentage of agreement: number of units scored identically/number of units scored × 100.

After transcriptionists achieved 80% reliability with the master videos, ongoing interrater reliability was calculated on 20% of the videos used in data analysis. Reliability was as follows: utterance segmentation M = 86%, total complete and intelligible utterances M = 96%, total number of words/line M = 87%, individual words M = 94%, total number of morphemes M = 81%, and individual morphemes M = 95%. Reliability among all English transcriptionists was M = 90% across categories.

Two senior undergraduate students who also participated in the coding and transcription of the English language videos transcribed the videos of the Spanish-speaking dyads using the standard Spanish conventions for SALT (Miller & Iglesias, 2016). These students were studying Spanish as a second language and as part of their degree programme. The two students trained on English videos until they were over 80% reliable across categories. Students then trained on three Spanish videos not used in the data analysis by transcribing videos independently and then developing a consensus transcript. Once transcribers were trained, they independently transcribed videos used in data analysis.

Ongoing Spanish transcription reliability was as follows: utterance segmentation M = 86%, total complete and intelligible utterances M = 98%, total number of words/line M = 92%, and individual words M = 96%. A native Spanish-speaking undergraduate speech-language-hearing sciences student performed a final review of each transcript with its corresponding video, editing as needed.

The SALT-derived measures included NTW, NDW, NumUtt, MTLUtt, MLU-M (English), and MLU-W (Spanish). In our analyses, we made a distinction between MLU-M for English transcripts and MLU-W for Spanish transcripts as reliable measures of sentence complexity for the respective languages. MLU-M is a calculation of the number of bound and free morphemes in a single utterance. Morphemes are structured differently in English and Spanish; therefore, we analysed the MLU-W for the Spanish samples. Gutiérrez-Clellen et al. (2000) discuss the complexities of counting morphemes in Spanish. Gender and number agreement must be considered morphemes in Spanish because of the importance of gender and number agreement in phrases such as "los gatos" (the cats) and "sus zapatos" (their shoes). Gutiérrez-Clellen et al. demonstrate the discrepancy of the morpheme count among researchers. As reported in Gutiérrez-Clellen et al. (2000), Linares and Sanders (1977) counted los perros as having six morphemes, but

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Garci (1978) counted only five morphemes in the phrase. The discrepancy among professionals and the vast difference between the morpheme count in English (3) and Spanish (5-6) makes it impossible to compare English MLU-M directly with Spanish MLU-M.

Behavioural coding, reliability, and measures

Once the transcribers completed transcription, trained coders coded each 6-minute video for communicative behaviour acts at the utterance level using Behavioral Observation Research Interactive Software (BORIS; Friard & Gamba, 2016), a free open-source event-logging software for video and audio coding. They coded communicative acts using a multitiered coding system adapted from Warren et al. (2010), which is a robust coding system that addresses multiple domains and has been used in previously peer-reviewed publications (Brady et al., 2014; Sterling & Warren, 2013). Coders assigned parental communicative acts codes corresponding to both levels of attention (e.g., Maintain, Introduce, Redirect) and function (e.g., Request for Verbal Reply, Behavioural Comply, Comment). Within each tier (i.e., Attention Level, Function), behavioural codes were mutually exclusive and codes were exhaustive across tiers. Table 2 presents the behavioural code and definitions included in this study. Complete coding details can be obtained by contacting the authors.

Advanced research assistants who created master videos trained coders on the multitiered coding system. Coders trained on three to five videos until they were at least 80% reliable on attention-level and function-level code categories. Coders used a manual that included detailed explanations of each code and discussed any specific questions in weekly lab meetings. All coders were blind to study details, group assignment, and whether the interaction was recorded pre- or post-intervention. We compared 20% of the videos for ongoing reliability. Two students, undergraduate and Master level, coded the videos of the English-speaking dyads. Their reliability for attention-level codes and function-level codes was $M = 88\%$ and $M = 87\%$, respectively. The same two students who transcribed the samples of the Spanish-speaking dyads also coded the videos for communicative acts. Their reliability for attention-level codes and function-level codes was $M = 91\%$ and $M = 87\%$, respectively.

Statistical analyses

We analysed Study Questions 1 and 3 using Welch's two-sample *t*-tests, which compared the gain (change) scores of the Full Sample (English and Spanish) LENA Start™ participants ($n = 28$) to the Comparison participants (English and Spanish $n = 16$). We used Welch's two-sample *t*-tests because it allows for a more stable analysis of uneven sample sizes and variance on gain scores than other tests. We calculated participant gain scores by subtracting pre-intervention counts of each variable from post-intervention counts. We also analysed group differences between treatment and comparison groups, controlling for pre-test performance, using ANCOVAs, a more sensitive measure to variance within the dependent variable.

To address Study Questions 2 and 4, we conducted within-group Welch's paired *t*-tests for the English-speaking LENA Start™ group ($n = 19$), the English-speaking Comparison group ($n = 15$), and the Spanish-speaking LENA Start™ group ($n = 9$). Because MLU was calculated differently for the English and Spanish groups, it is not reported for the Full

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Sample. We collected data from only one parent–child dyad in the Spanish-speaking Comparison group; therefore, we only present descriptive statistics for this group. Due to the limited sample size, we did not conduct analyses between the Spanish-speaking intervention group (n = 9) and the Spanish-speaking Comparison group (n=1). Instead, we include the descriptive data in Table 1 to allow readers to better understand the demographics of the Full Intervention and Full Comparison groups.

Given our directional hypotheses, we used one-tailed t-tests. To calculate effect sizes for the t-tests, we converted Cohen's d to hedges g values, reducing bias of the standardised mean in the small sample size. Hedges g is more appropriate for sample sizes of less than 20. We interpreted effect sizes using the following guidelines from Cohen (1988): small effect: g = .20–.49, medium effect: g = .50–.79, and large effect: $g \geq .80$. Effect sizes associated with the ANCOVA results are eta squared (η^2): small η^2 = .01–.059, medium η^2 = .06–.139, and large $\eta^2 \geq .14$. We also report 95% confidence intervals (CIs). Note that when zero falls within a CI, we cannot claim that there is a true difference between the groups being compared. We used R Studio Version 4.0.3 to complete the analyses (R Studio Team, 2020).

Results

Research Question 1: Comparison of LENA Start™ and Comparison Groups on SALT Measures

To evaluate group differences based on adult language change based on SALT measures in the Full Sample LENA Start™ and Comparison groups, we used t-tests and ANCOVAs. Results of the Welch's two-sample t-tests revealed no statistical difference between the LENA Start™ and Comparison groups for adult language change from pre- to post-intervention observations. Table 3 contains the gain scores and associated p- and g-values for each of the SALT variables for the Full Sample. All t-tests yielded p-values greater than .11 and hedge's g effect sizes less than 0.41, which are considered small. Boxplots in Figure 1 visually represent the large overall performance range across measures and the overlap of the interquartile range among all measures, indicating no meaningful differences between groups. Some standard deviations (SD) are larger than the means (e.g., NTW, NDW), which indicates that the data are highly dispersed and therefore it is difficult to determine the true mean. Results from the ANCOVAs indicated that there was no significant difference between the LENA Start™ and Comparison groups when controlling for pre-test values on all measures (see Table 4 for ANCOVA SALT measure results). Results of the ANCOVA were associated with small effect sizes (η^2s < .021).

Research Question 2: Within-Group Comparisons of Pre- and Post-intervention SALT Measures

To evaluate within-group differences in adult language for the English- and Spanish-speaking participants in the LENA Start™ programme based on SALT measures, we used t-tests and ANCOVAs. Table 5 contains each study groups' pre- and post-intervention values, p-values and g-values for each SALT measure. Table 5 first shows within-group analyses of the English-speaking LENA Start™ group followed by the within-group analysis for the English-speaking Comparison group. Four of the five measures of the treatment group were associated with p-values greater than .05 and effect sizes less than

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Table 1. Participant characteristics

	English-speaking		Spanish-speaking	
	LENA Start™ (*n* = 19)	Comparison (*n* = 15)	LENA Start™ (*n* = 9)	Comparison (*n* = 1)
Child Age in Months Mean (SD)	16.05 (5.31)	23.80 (8.66)	21.44 (9.41)	22
Child Sex Assigned at Birth				
Male (%)	12 (63%)	6 (40%)	4 (44%)	1 (100%)
Female (%)	7 (37%)	9 (60%)	5 (56%)	0
Child Race/Ethnicity				
Asian/Asian-American (%)	1 (5%)	0 (0%)	0 (0%)	0 (0%)
Black/African American (%)	1 (5%)	0 (0%)	0 (0%)	0 (0%)
Hispanic/Latino (%)	0 (0%)	0 (0%)	9 (100%)	1 (100%)
White/Caucasian (%)	16 (85%)	15 (100%)	0 (0%)	0 (0%)
Multi-Ethnic (%)	1 (5%)	0 (0%)	0 (0%)	0 (0%)
Parent Age in Years Mean (SD)	30.53 (5.61)	31.14 (5.02)	34.22 (5.09)	19
Home Language				
English (%)	16 (85%)	14 (93%)	0 (0%)	0 (0%)
Spanish (%)	0 (0%)	0 (0%)	5 (56%)	0 (0%)
Bilingual SP&EN (%)	0 (0%)	0 (0%)	3 (33%)	1 (100%)
Other Bilingual (%)	3 (15%)	1 (7%)	1 (11%)	0 (0%)
Family Size (≥ 4) (%)	15 (79%)	13 (87%)	4 (44%)	1 (100%)

Note. Home language = language adult uses when talking to key child, per parent report. Other Bilingual = Mandarin, German, and American Sign Language and English; SP&EN = Spanish and English.

.22. MLU-M was associated with a *p*-value of .04 and a near medium effect size ($g_z = 0.49$). English-speaking Comparison group results were associated with *p*-values greater than .05 and hedge's *g* effect sizes less than .28.

Within-group analyses of Spanish-speaking LENA Start™ participants yielded three tests with *p*-values less than .05 (NTW, NDW, and NumUtt). Additionally, NumUtt was associated with a medium effect size ($g_z = 0.56$), while NTW, NDW, and MLU-W were all associated with large effect sizes ($g_z = 0.95$, 1.34, and 1.26, respectively).

Research Question 3: Comparison of LENA Start™ and Comparison Groups on Communication behaviour measures

To evaluate between-group differences in adult language change based on coded communicative behaviour measures in the Full Sample LENA Start™ and Comparison groups,

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Table 2. Parent communicative behaviours codes

Category & Behaviour	Definition
Level of Attention	
Maintain	Parent references toy, behaviour, or emotional state of child.
Redirect	Child is actively involved physically and/or visually in play with an object different from one that the parent presents/ references. Parent is "shifting gears" by referencing a new object when the child is actively attending to another object.
Introduce	Occurs when parent presents either themself or a new object/ activity at a time when a child is not actively attending to the item being presented. Child is not looking or touching objects for 1+ second/s. *Default: This is the default code when it is impossible to tell to what the child is attending.*
Reading	Parent reads verbatim (with reasonable certainty) without comment, further explanation/description, or question directed toward child or self.
Function	
Request for behavioural comply	Parent's intent with the communicative act is to provide directives to which the child can comply behaviourally. Although questions such as "do you want to sit down" are phrased like a request for verbal reply, the intent is for the child to sit down.
Request for verbal reply	Parent asks questions or gives a model then pauses 3 seconds or more. The parent asks how the child's day was or begins a conversation.
Comment	Talking about what the child or parent can see, hear, smell, taste, touch. Praise or phrases in reaction to something the child has done including "whoops" or "uh oh."

we used *t*-tests and ANCOVAs. Table 6 contains the gain scores and associated *p*- and *g*-values for each of the communication behaviour variables for the Full Sample. All *t*-tests were associated with *p*-values greater than .20 and effect sizes less than 0.40, which are

Table 3 *T-test and effect size results for between-group comparison of Full Sample gain scores based on SALT measures*

	LENA Start™ (n = 28)		Comparison (n = 16)					
Variable	*M*	SD	*M*	SD	95% CI	*t*(42)	*p*	Hedge's *g*
NTW	57.82	100.82	21.25	90.73	[−23.6, 96.8]	1.235	.113	0.37
NDW	18.25	37.13	6.75	25.21	[−7.6 30.6]	1.219	.115	0.34
NumUtt	8.25	27.73	4.94	21.73	[−12.0, 18.6]	0.439	.332	0.41
MTLUtt	−0.87	14.22	1.06	6.83	[−8.4, 4.5]	−0.606	.274	−0.16

Note. M = mean, SD = standard deviation, CI = confidence interval, NTW = number of total words, NDW = number of different words, NumUtt = total number of utterances, MTLUtt = mean turn length of utterance.

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

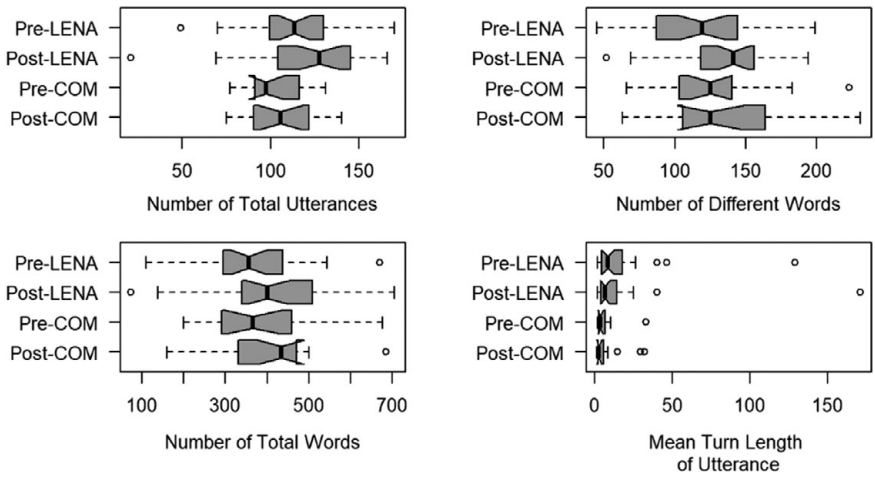

Figure 1. *Boxplots of SALT Pre- and Post-intervention measures for Full Sample LENA Start™ (LENA) and Comparison (COM) Groups*

considered small. ANCOVA results (Table 7) indicate statistically significant results based on Redirect (p = .011, η^2 = .147) and Introduce (p = .030, η^2 = .110) variables. Maintaining (η^2 = .010), Verbal Reply (η^2 = .020), Reading (η^2 = .013), and Behaviour Compliance (η^2 = .078) were all associated with small or medium effect sizes, though they are not statistically significant.

Research question 4: within-group comparisons of pre- and post-intervention Communication behaviour measures

To evaluate within-group differences in adult language for the English- or Spanish-speaking participants in the LENA Start™ group based on communicative behaviour measures, we used *t*-tests and ANCOVAs. Table 8 contains each study groups' pre- and post-intervention values, *p*-values, and *g*-values for each of the communication

Table 4 *ANCOVA results for English-speaking participants controlling for pretest scores*

	LENA		Comparison				
Variable	Pre	Post	Pre	Post	F	*p*	η^{2e}
NTW	350.21 (129.93)	408.04 (130.40)	383.62 (128.06)	404.88 (120.39)	.863	.358	.021*
NDW	118.61 (38.33)	136.86 (32.50)	128.38 (37.78)	135.12 (40.53)	.605	.441	.015*
NumUtt	112.79 (26.31)	121.04 (32.26)	101.88 (16.54)	106.81 (19.30)	.872	.356	.021*
MTLUtt	16.35 (24.72)	15.48 (31.71)	6.20 (7.58)	7.27 (9.74)	.790	.379	.019*

Note. NTW = number of total words, NDW = number of different words, NumUtt = total number of utterances, MTLUtt = mean turn length of utterance. Effect sizes:
*small (.01 – .059), **medium (.06 – .139), ***large (≥.14).

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Table 5 *T-test and effect size results for within-group analyses of SALT measures*

	PRE-intervention		POST-intervention					
Variable	*M*	*SD*	*M*	*SD*	95% CI	*t*	*p*	hedge's *g*
English-speaking LENA Start™ (*n* = 19)								
NTW	401.63	111.99	430.21	131.44	[−17.7, 74.9]	1.297	.106	0.22
NDW	135.84	30.17	138.00	30.64	[−12.4 16.7]	0.311	.379	0.07
MLU-M	3.75	0.50	4.01	0.53	[0.0, 0.6]	1.838	.042ŧ	0.49
NumUtt	115.26	24.08	117.89	32.22	[−9.6, 14.9]	0.452	.328	0.09
MTLUtt	12.30	10.19	9.22	9.18	[−9.1, 3.0]	−1.070	.150	−0.30
English-speaking comparison (*n* = 15)								
NTW	393.67	125.87	421.20	104.69	[−22.4, 77.5]	1.182	.129	0.22
NDW	130.93	37.65	139.93	36.93	[− 4.5, 22.5]	1.43	.088	0.23
MLU-M	4.09	0.90	4.28	0.78	[− 0.2, 0.5]	1.215	.123	0.22
NumUtt	102.67	16.80	108.07	19.29	[− 7.0, 17.8]	0.933	.183	0.28
MTLUtt	6.38	7.81	7.60	9.99	[− 2.7, 5.1]	0.671	.257	0.12
Spanish-speaking LENA Start™ (*n* = 9)								
NTW	241.67	96.06	361.22	121.99	[54.2, 184.9]	4.221	.002ŧ	0.95***
NDW	82.22	26.74	134.44	37.97	[32.1, 72.4]	5.983	< .001ŧ	1.34***
MLU-W	2.33	0.61	2.51	0.75	[−0.4, 0.8]	0.726	.245	1.26***
NumUtt	107.56	31.39	127.67	33.21	[−3.1, 43.3]	1.998	.041ŧ	0.56**
MTLUtt	24.92	41.31	28.70	53.98	[−9.4, 17.0]	0.662	.264	0.05
Spanish-speaking comparison Group (*n* = 1)								
NTW	233.00	na	160.00	na	na	na	na	na
NDW	90.00	na	63.00	na	na	na	na	na
MLU-W	1.78	na	2.42	na	na	na	na	na
NumUtt	90.00	na	88.00	na	na	na	na	na
MTLUtt	3.63	na	2.30	na	na	na	na	na

Note. M = mean, SD = standard deviation, CI = confidence interval, NTW = number of total words, NDW = number of different words, MLU-M = mean length of utterance by morpheme, NumUtt = total number of utterances, MTLUtt = mean turn length of utterance.
ŧ*p* > .05.
**medium effect size, ***large effect size.

behaviour measures. For all measures, within-group analyses for both the English-speaking LENA Start™ group and the English-speaking Comparison group yielded *p*-values greater than .05 that were associated with small effect sizes. For the Spanish-speaking LENA Start™ group, the pre-/post-intervention comparison for Maintaining had a *p*-value less than .05 and was associated with a large effect size (g_z = 0.79). Verbal Reply had a *p*-value less than .05 and was also associated with a large effect size (g_z = 0.84).

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Table 6 *T-test and Hedge's g effect size results for between-group comparison of Full Sample gain scores based on communication behaviour measures*

	LENA Start™ (*n* = 28)		Comparison (*n* = 16)					
Variable	*M*	*SD*	*M*	*SD*	95% CI	*t*(42)	*p*	Hedge's *g*
Maintain	4.86	39.27	4.81	18.06	[−21.0, 21.1]	0.004	.499	>0.01
Redirect	1.64	7.46	−0.88	3.01	[−1.4, 6.5]	1.286	.103	0.40
Comment	−2.96	26.08	3.75	11.02	[−20.6, 7.2]	−0.977	.167	−0.31
Verbal reply	7.43	19.40	2.44	16.97	[−6.8, 16.7]	0.858	.198	0.27
Introduce	−1.07	3.63	−0.31	4.90	[−3.4, 1.9]	−0.586	.281	−0.18
Reading	−0.04	5.90	−0.63	10.40	[−4.3, 5.5]	0.241	.401	0.07
Behaviour Comply	1.00	15.71	−2.63	7.14	[−4.8, 12.0]	0.870	.195	0.27

Note. M = mean, SD = standard deviation, CI = confidence interval.

Table 7 *ANCOVA and eta squared effect size results controlling for pre-test of Full Sample communication behaviour measures*

	LENA Start™ (*n* = 28)		Comparison (*n* = 16)				
Variable	Pre-Mean (SD)	Post-Mean (SD	Pre-Mean (SD)	Post-Mean (SD)	F	*P*	η^2
Maintain	108.82 (26.02)	113.68 (40.12)	95.44 (20.43)	100.25 (15.51)	.415	.523	.010*
Redirect	3.89 (4.66)	5.54 (6.02)	2.06 (2.89)	1.19 (1.60)	7.060	.011[ŧ]	.147***
Comment	58.57 (17.19)	55.61 (21.10	48.81 (12.16)	52.56 (13.60)	.015	.903	> .001
Verbal Reply	39.86 (20.08)	47.29 (21.16)	40.44 (15.33)	42.88 (14.22)	.829	.368	.020*
Introduce	3.07 (3.09)	2.00 (2.04)	3.81 (4.15)	3.50 (2.22)	5.060	.030[ŧ]	.110**
Reading	2.00 (4.20)	1.96 (4.54)	4.50 (7.91)	3.88 (8.45)	.555	.461	.013*
Behaviour Comply	17.39 (16.54)	18.39 (15.40)	12.00 (7.43)	9.38 (4.50)	3.465	.070	.078**

[ŧ]*p* > .05, Effect sizes.
*small (.01–.059), **medium (.06–.139), ***large (≥ .14) (Cohen 1988).

Discussion

The purpose of this study was to conduct a follow-up evaluation of relatively distal outcomes after parents participated in a parent-implemented intervention (i.e., LENA Start™) to modify their child's language environment. Multiple studies have been conducted on the use of LENA technology and how feedback from recordings of the child's language environment can influence change (e.g., Magimairaj et al., 2022). Additionally, both Beecher and Van Pay (2020) and Elmquist et al. (2021) evaluated the use of LENA technology in conjunction with the LENA Start™ programme. The current study extends previous investigations by evaluating the generalisability of the LENA Start™ programme using measures not directly associated with LENA and in a context outside of the home.

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Table 8 *T-test and effect size results for within-group analyses of communication behaviour measures*

	PRE-Intervention		POST-Intervention					
Variable	*M*	*SD*	*M*	*SD*	95% CI	*t*	*p*	Hedge's *g*
English-speaking LENA Start™ (*n* = 19)								
Maintain	108.79	23.14	103.16	39.58	[−24.6, 13.4]	−0.622	.271	−0.15
Redirect	3.32	4.26	5.74	5.62	[−0.6, 5.5]	1.670	.056	0.44
Comment	59.68	17.35	54.42	21.65	[−18.4, 7.9]	−0.839	.206	−0.24
Verbal reply	42.79	19.46	44.95	21.33	[−4.3, 8.6]	0.706	.245	0.09
Introduce	3.89	3.41	2.42	2.32	[−3.6, 0.6]	−1.494	.077	−0.46
Reading	2.32	4.61	2.00	4.48	[−3.3, 2.6]	−0.226	.412	−0.06
Behaviour Comply	13.58	7.86	11.95	8.91	[−4.7, 1.4]	−1.118	.139	−0.17
English-speaking comparison (*n* = 15)								
Maintain	95.67	21.12	101.13	15.63	[−4.8, 15.7]	1.144	.136	0.27
Redirect	1.87	2.88	1.27	1.62	[−2.2, 1.0]	−0.802	.218	−0.24
Comment	48.6	12.55	51.53	13.42	[−3.1, 9.0]	1.043	.158	0.21
Verbal reply	41.6	15.12	44.73	12.55	[−6.5, 12.7]	0.700	.248	0.21
Introduce	4	4.23	3.67	2.19	[−3.1, 2.5]	−0.255	.402	−0.09
Reading	4.8	8.09	4.13	8.68	[−6.6, 5.3]	−0.240	.406	−0.08
Behaviour Comply	11.27	7.07	9.67	4.50	[−5.0, 1.8]	−1.023	.162	−0.24
Spanish-speaking LENA Start™ (*n* = 9)								
Maintain	108.9	32.87	135.9	33.00	[4.1, 49.9]	2.716	.013ŧ	0.79**
Redirect	5.1	5.487	5.1	7.13	[−7.4, 7.4]	0	.500	No change
Comment	56.2	17.64	58.1	20.93	[−16.6, 20.3]	0.236	.410	0.09
Verbal reply	33.7	21.11	52.2	21.13	[−1.3, 38.4]	2.159	.032ŧ	0.84***
Introduce	1.3	1	1.1	0.78	[−1.2, 0.8]	−0.513	.311	−0.24
Reading	1.3	3.317	1.8	4.94	[−3.9, 5.0]	0.288	.391	0.13
Behaviour Comply	25.4	26.01	32	17.76	[−13.7, 26.8]	0.748	.238	0.28
Spanish-speaking Comparison (*n* = 1)								
Maintain	92	na	87	na	na	na	na	na
Redirect	5	na	0	na	na	na	na	na
Comment	52	na	68	na	na	na	na	na
Verbal Reply	23	na	15	na	na	na	na	na
Introduce	1	na	1	na	na	na	na	na
Reading	0	na	0	na	na	na	na	na
Behaviour Comply	23	na	5	na	na	na	na	na

Note. M = mean, SD = standard deviation, CI = confidence interval.
ŧ*p* > .05.
**medium effect size, ***large effect size.

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

We collected parent–child free play samples from each parent–child dyad prior to and after intervention from which we derived distal language and communicative behaviour measures. Both sets of measures were drawn from a situation different from the intervention context (e.g., researcher-provided toys and books). Results of SALT analyses for the Spanish-speaking group demonstrated change for intervention participants in NTW, NDW, and MLU-W and were associated with large effects. NumUtt increased by a smaller margin. Examination of Table 5 reveals that the mean NTW and NDW pre-intervention values of the Spanish-speaking parents were considerably lower than the English-speaking parents' values but similar to post-intervention values. Thus, the Spanish-speaking parents had more room for change and did indeed make significant gains on these two measures.

ANCOVA and *t*-test analyses did not reveal statistically significant results across the SALT measures in analyses between the Full Sample LENA Start™ and Comparison groups. ANCOVA and *t*-test results revealed small gains not associated with statistical significance for NTW, NDW, NumUtt, and MTLUtt. The lack of statistical significance is likely due to the large range that can be seen in the raw numbers among participants pre- and post-intervention use of language measures. These ranges are reflected in the standard deviations associated with means (see Tables 3 and 5) and are likely due to the relatively small sample size. A larger sample size allows for nuance within statistical analysis in heterogeneous samples. Additionally, we cannot rule out that some of the participants elected to participate in the intervention group and therefore may have been more eager to use the skills being taught.

Another potential explanation for the variability in results could be that the researchers did not situate the parents into the mindset of LENA Start™ intervention. Although the parents may have benefitted from explicit guidance or reminders of the Talking Tips learned in the intervention from the researchers, the assessment was designed to serve as a rigorous evaluation of the generalisation of parents' communication-facilitative skills. We wanted to determine if parents can translate what they learned in the intervention to settings in which they did not receive specific guidance, similar to how they might interact with their children in real-life settings.

Communicative behaviour measure results were different from SALT language measures; most differences were found in the results of the ANCOVA analyses. No *t*-tests resulted in statistically significant outcomes based on the mean change score between the LENA Start™ and the Comparison groups and all analyses were associated with small effect sizes (gs $\leq .041$). Comment and Introduce effect sizes were associated with the Comparison group having better results than the LENA Start™ group. This could be attributable to the large differences seen in the standard deviations compared to the means. Within-group comparisons of parents' ability to maintain their child's language focus and request verbal replies yielded significant changes, which were associated with medium and large effect sizes, respectively.

ANCOVA results of communicative behaviour measures revealed multiple statistically significant outcomes and large effect sizes. Because ANCOVA is a more sensitive measure to variance than *t*-tests, results reveal positive outcomes of LENA Start™ intervention on distal measures. Parents significantly decreased (i.e., positive outcome) their use of redirecting a child to object/activity and introducing a new object/activity ($\eta^2 = .147$ and .110, respectively). Additionally, parents improved in maintaining their child's communication topic ($\eta^2 = .010$), requesting the child to say something ($\eta^2 = .020$), and requesting the child to do something ($\eta^2 = .078$).

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Evaluating an intervention using measures that are highly related and proximal to the intervention is meaningful to determine the direct impact of the intervention. Distal measures, however, evaluate the broader impact of the intervention. While initial and more proximal intervention effects are likely a necessary condition for efforts focused on clinical utility, they may not be sufficient. An intervention that builds language skills within one specific context, but that does not transfer to other communicative situations, is insufficient when the goal of the language intervention is to improve daily communication experiences.

One unique feature of LENA Start™ is the use of the LENA technology to acquire weekly logs of AWC, CTC, and CVC. Interventionists coach parents on strategies to increase these measures to enhance their child's language environment. While Elmquist et al. (2021) and Beecher and Van Pay (2019; 2020) reported positive LENA Start™ outcomes based on these measures, our study focused on distal measures that researchers have not previously evaluated. Our preliminary findings based on these distal measures suggest that LENA Start™ outcomes outside intervention contexts may not yet be robust.

One reason for the differences in outcomes based on proximal and distal measures may be the sampling context. The LENA measures are based on day-long samples (up to 16 hours), typically recording the child at home or in other familiar contexts. In contrast, the free play samples were short (6 minutes) and recorded in an early childhood centre less familiar to the child and parent. Thus, 6 minutes may not be sufficient time for the child to feel comfortable to readily communicate and for the parent to have enough opportunities to demonstrate changes in language use and communicative behaviours. However, Heilman et al. (2010) evaluated the results of 1-minute, 3-minute, and 7-minute language samples for children using conversational and narrative contexts. Researchers found no significant differences and negligible effect sizes for differences among these lengths of language samples. Cronbach's alpha analyses revealed that 3-minute samples were more reliable than the 1-minute samples but did not substantially vary from the 7-minute sample (Heilman et al., 2010). More research is needed to identify valid and reliable contexts from which to derive distal measures to evaluate communication interventions for young children.

It may also be the case that the LENA device itself may serve as a trigger for the parent to focus on increasing their language input and promoting turn-taking. In this study, we provided parents with a set of toys and books and asked them to play with their child as they would at home. We then video-recorded the interaction. To rigorously evaluate the impact of the intervention, we did not remind parents to use LENA Start™ strategies. Thus, it is possible that when the child is not wearing the device, parents may forget to use the strategies taught in the LENA Start™ intervention.

Another plausible explanation for outcome differences between this study and previous examinations of LENA Start™ may be due to variation in the specific measures derived from the samples. The gross measure of AWC does not compare easily with the more specified SALT and communicative behaviour measures analysed in this study. For example, Redirect may be used as a behaviour management tool by parents but Maintain may be a more appropriate behaviour management tool in a free play setting because of the joint attention parent and child share. When calculating AWC, parents are "credited" for every word they use, regardless of appropriateness. The nuances of individual child language needs are difficult to compare to gross measures of simple increase of language use.

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Study Limitations and Future Research Directions

Limitations of this study mark a path for future research. One limitation was the small sample size, particularly, the size of the Spanish-speaking sample. Despite the small sample sizes, analyses were associated with medium and large effect sizes, particularly for the Spanish-speaking treatment group and communicative behaviour measures. This suggests that the intervention may have positive gains on distal measures; however, further evaluation with larger sample sizes is needed to better understand the generalised intervention effects. It is also possible that differential effects, and thus, sometimes larger intervention effects, are associated with one or more person-by-treatment interactions. Restricted resources limited the scope of this study; however, preliminary results of this small sample who have a wide range of individual differences have positive trending outcomes. A larger and more narrowly defined sample may reveal important findings in future studies and long-term follow-ups. For instance, Spanish-speaking families' variables were lower at pre-intervention, giving them more opportunity for change. Perhaps follow-up investigation of families specifically selected for initial measure status would help better understand this possible relation.

Based on the results of previous studies, LENA Start™ can make a positive difference in proximal measures. Our results suggest that there may also be a positive effect on distal measures. Future studies should allow for a larger sample size and randomisation. Continued study of distal outcome measures would benefit parent-implemented interventions for young children. The use of distal measures is imperative to assess the generalisability of intervention effects. Researchers must continue to use distal measures for a broader evaluation of language intervention.

In conclusion, we find that this cohort of participants had limited success in improving their language skills, based on more distal measures, after a LENA Start™ intervention. Exception to this is that for Spanish-speaking parents, results suggested some promising positive intervention effects. Thus, further investigation of LENA Start™ is warranted.

Acknowledgements. The authors thank administrators and staff of St. Paul Public Schools, Anoka-Hennepin Schools, and other participating districts, as well as the parents and children who participated in this work. We are also grateful for assistance from Anne Larson, Amanda Kriese, Erin Lease, and Gerri Fischer in launching this effort, and ongoing support from Jill Gilkerson, Traci Martin, Kim Coulter, and Beth Dresser from the LENA Research Foundation. Research reported here was supported in part with funds from the Bush Foundation in partnership with ThinkSmall and Generation Next, the University of Minnesota, and LENA Research Foundation. All work was designed and is the sole responsibility of the authors with no advance reviews from funders, and no endorsements by these funders should be implied.

References

Beecher, C. C., & **Van Pay, C. K**. (2019). Small talk: A community research collaboration to increase parental provision of language to children. *Child & Youth Care Forum*, **50**(1), 13–38. https://doi.org/10.1007/s10566-019-09507-7

Beecher, C. C., & **Van Pay, C. K**. (2020). Investigation of the effectiveness of a community-based parent education program to engage families in increasing language interactions with their children. *Early Childhood Research Quarterly*, **53**, 453–463. https://doi.org/10.1016/j.ecresq.2020.04.001

Brady, N., **Warren, S. F.**, **Fleming, K.**, **Keller, J.**, & **Sterling, A.** (2014). Effect of sustained maternal responsivity on later vocabulary development in children with fragile X syndrome.

Bronfenbrenner, U., & **Ceci, S. J.** (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review*, **101**(4), 568–586. https://doi.org/10.1037/0033-295X.101.4.568

Carle, E. (1969). The Very Hungry Caterpillar. World Publishing Company. ISBN 688 0-399-22690-7. 689
Carle, E. (2002). *La oruga muy hambrienta.* Penguin Young Readers. ISBN 039923960X.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
DK. (2004). *Buenos Días, Bebé!/Good Morning, Baby!* DK Children; Bilingual edition. ISBN 697 0756604370. 698
Dorling Kindersley, Inc. (2008). *Touch and feel: Baby animals*. DK Publisher.
Elmquist, M., Finestack, L. H., Kriese, A., Lease, E. M., & **McConnell, S. R**. (2021). Parent education to improve early language development: A preliminary evaluation of LENA Start™. *Journal of Child Language*, **48**(4), 670–698. https://doi.org/10.1017/s0305000920000458
Finestack, L. H., Elmquist, M., Kuchler, K., Ford, A. L. B., Çakır-Dilek, B., Riegelman, A., Brown, S. J., & **Marsalis, S**. (2022). Caregiver-implemented communication interventions for children identified as having language Impairment 0 Through 3 Years of age: A scoping review. *Journal of Speech, Language, Hearing Research*. https://doi.org/10.1044/2022_JSLHR-21-00543
Friard, O. P., & **Gamba, M**. (2016). Behavioral observation research interactive software (BORIS). 2016. Available from: https://iris.unito.it/handle/2318/1589424#.X7x7grMRXIU. Accessed 23.11.20).
Ford, A. L. B., Elmquist, M., Merbler, A. M., Kriese, A., Will, K. K., & **McConnell, S. R**. (2020). Toward an ecobehavioral model of early language development. *Early Childhood Research Quarterly*, **50**(1), 246–258. https://doi.org/10.1016/j.ecresq.2018.11.004
Greenwood, C. R., Thiemann-Bourque, K., Walker, D., Buzhardt, J., & **Gilkerson, J**. (2010). Assessing children's home language environments using automatic speech recognition technology. *Communication Disorders Quarterly*, **32**(2), 83–92. https://doi.org/10.1177/1525740110367826
Gutiérrez-Clellen, V. F., Restrepo, M. A., Bedore, L., Peña, E., & **Anderson, R**. (2000). Language sample analysis in Spanish-speaking children. *Language, Speech, and Hearing Services in Schools*, **31**(1), 88–98. https://doi.org/10.1044/0161-1461.3101.88
Heidlage, J. K., Cunningham, J. E., Kaiser, A. P., Trivette, C. M., Barton, E. E., Frey, J. R., & **Roberts, M. Y.** (2019). The effects of parent-implemented language interventions on child linguistic outcomes: A meta-analysis. https://doi.org/10.1016/j.ecresq.2018.12.006
Heilman, J., Nockerts, A., & **Miller J. F**. (2010). Language sampling: Does the length of the transcript matter? *Language, Speech, and Hearing Services in Schools*, **41**, 393–404. https://doi.org/10.1044/0161-1461(2009/09-0023)
Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., Yust, P. K., & **Suma, K**. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, **26**(7), 1071–1083. https://doi.org/10.1177/0956797615581493
Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, **26**(1), 55–88. https://doi.org/10.1016/j.dr.2005.11.002
LENAEarlyTalk, via. (2022, June 10). *Developmental snapshot.* LENA. Retrieved July 29, 2022, from https://www.lena.org/developmental-snapshot/#:~:text=The%20Developmental%20Snapshot%20is%20widely,to%20accelerate%20children's%20language%20development.
Linares-Orama, N., & **Sanders, L. J.** (1977). Evaluation of syntax in three-year-old Spanish-speaking Puerto Rican children. *Journal of Speech and Hearing Research*, **20**(2), 350–357. https://doi.org/10.1044/jshr.2002.350
Luze, G. J., Linebarger, D. L., Greenwood, C. R., Carta, J. J., Walker, D., Leitschuh, C., & **Atwater, J. B.** (2001). Developing a general outcome measure of growth in the expressive communication of infants and toddlers. *School Psychology Review*, **30**(3), 383–406.
Magimairaj, B. M., Nagaraj, N. K., Caballero, A., Muñoz, K. F., & **White, K. R**. (2022). A systematic review of the effects of LENA-based feedback on parent-child language interactions in families with young children. *The Journal of Early Hearing Detection and Intervention*, **7**(3), 47–60.
Miller, J., & **Iglesias, A**. (2016). Systematic analysis of language transcripts (SALT), Research Version 2016 [Computer Software]. Middleton, WI: SALT Software, LLC.
Noble, C., Cameron-Faulkner, T., Jessop, A., Coates, A., Sawyer, H., Taylor-Ims, R., & **Rowland, C. F.** (2020). The impact of interactive shared book reading on children's language skills: A randomized controlled trial. *Journal of Speech, Language, and Hearing Research*, **63**, 1878–1897. https://doi.org/10.1044/2020_JSLHR-19-00288

Richards, J. A., **Xu, D.**, **Gilkerson, J.**, **Yapanel, U.**, **Gray, S.**, & **Paul, T**. (2017). Automated assessment of child vocalization development using LENA. *Journal of Speech, Language, and Hearing Research*, **60**(7), 2047–2063. https://doi.org/10.1044/2017_JSLHR-L-16-0157

Roberts, M. Y., **Curtis, P. R.**, **Sone, B. J.**, **Hampton, L. H.** (2019). Association of parent training with child language development: A systematic review and meta-analysis. *JAMA Pediatrics*, **173**(7), 671–680. https://doi.org/10.1001/jamapediatrics.2019.1197

Rstudio Team (2020). R Studio (Version 4.0.3 [Computer Software]. The R Foundation for Statistical Computing. https://www.rstudio.com/products/rstudio/download/#download

Sameroff, A. J. (2009). *The transactional model of development how children and contexts shape each other*. American Psychological Association.

Sandbank, M., **Chow, J.**, **Bottema-Beutel, K.**, & **Woynaroski, T.** (2021). Evaluating evidence-based practice in light of the boundedness and proximity of outcomes: Capturing the scope of change. *Autism Research*, **14**(8), 1536–1542.

Sterling, A., & **Warren, S. F.** (2013). Maternal responsivity in mothers of young children with Down syndrome. *Developmental Neurorehabilitation*, **17**(5), 306–317. https://doi.org/10.3109/17518423.2013.772671

Stokes, T. F., & **Baer, D. M**. (1977). An implicit technology of generalization 1. *Journal of Applied Behavior Analysis*, **10**(2), 349–367.

Suskind, D. L., **Leffel, K. R.**, **Graf, E.**, **Hernandez, M. W.**, **Gunderson, E. A.**, **Sapolich, S. G.**, **Suskind, E.**, **Leininger, L.**, **Goldin-Meadow, S.**, & **Levine, S. C**. (2016). A parent-directed language intervention for children of low socioeconomic status: A randomized controlled pilot study. *Journal of Child Language*, **43**(2), 366–406. https://doi.org/10.1017/s0305000915000033

Wan, M. W., **Green, J.**, **Elsabbagh, M.**, **Johnson, M.**, **Charman, T.**, & **Plummer, F.** (2012). Parent–infant interaction in infant siblings at risk of autism. *Research in Developmental Disabilities*, **33**(3), 924–932. https://doi.org/10.1016/j.ridd.2011.12.011

Warren, S. F., **Brady, N.**, **Sterling, A.**, **Fleming, K.**, & **Marquis, J.** (2010). Maternal responsivity predicts language development in young children with fragile X syndrome. *American journal on intellectual and developmental disabilities*, **115**(1), 54–75. https://doi.org/10.1352/1944-7558-115.1.54

Yoder, P. J., **Bottema-Beutel, K.**, **Woynaroski, T.**, **Chandrasekhar, R.**, & **Sandbank, M.** (2013). Social communication intervention effects vary by dependent variable type in preschoolers with autism spectrum disorders. *Evidence-Based Communication Assessment and Intervention*, **7**(4), 150–174. https://doi.org/10.1080/17489539.2014.917780

Appendix A

Description of LENA Start™ weekly sessions	
Topic	Description
Week 1: Introduction to LENA Start™	Overview of LENA Start™ programme and instructions on using LENA recording devices.
Week 2: LENA reports and the 14 talking tips	Introduces using LENA reports, as well as the 14 talking tips.
Week 3: Shared reading	Information on how to use shared reading to increase turn-taking and words, as well as practicing the 14 talking tips. Parents receive first LENA recording report
Week 4: Songs and rhymes	Information on how to incorporate songs, rhymes and fingerplay into parent–child interactions to increase turn-taking and words, as well as some more practice of the 14 talking tips.
Week 5: Talking tips practice and group report	Revisiting talking tips and going over group LENA report

https://doi.org/10.1017/S0305000924000606 Published online by Cambridge University Press

Description of LENA Start™ weekly sessions	
Topic	Description
Week 6: More about your baby's brain	Information on infant brain development and time for group discussion on experiences thus far
Week 7: Midpoint reflections	Session devoted to group reflections as well as practicing talking tips
Week 8: Math talk – movement	Information on incorporating movement words (e.g. fast, right, down etc.) into parent–child interactions to build math and language development
Week 9: Building brains by asking questions	Information on asking questions to build language
Week 10: Language of food	Information on incorporating language into mealtimes
Week 11: Math talk – space	Information on incorporating spatial words (e.g. on, under etc.) to build math and language development
Week 12: Out and about	Information on incorporating language while out and about
Week 13: Graduation day	Final group sharing time and review of talking tips

Cite this article: Kuchler, K., Elmquist, M., McConnell, S.R., & Finestack, L.H. (2025). An Evaluation of LENA Start™ Using Measures Derived from Parent–Child Interactions. *Journal of Child Language* 1–24, https://doi.org/10.1017/S0305000924000606