RESEARCH ARTICLE   **Open Access**

The education word gap emerges by 18 months: findings from an Australian prospective study

Mary E. Brushe[1,2*], John Lynch[2,3], Sheena Reilly[4], Edward Melhuish[5], Murthy N. Mittinty[2] and Sally A. Brinkman[1,2]

Abstract

Background: The idea of the '30 million word gap' suggests families from more socioeconomically advantaged backgrounds engage in more verbal interactions with their child than disadvantaged families. Initial findings from the Language in Little Ones (LiLO) study up to 12 months showed no word gap between maternal education groups.

Methods: Families with either high or low maternal education were purposively recruited into a five-year prospective study. We report results from the first three waves of LiLO when children were 6, 12 and 18 months old. Day-long audio recordings, obtained using the Language Environment Analysis software, provided counts of adult words spoken to the child, child vocalizations and conversational turns.

Results: By the time children were 18 months old all three measures of talk were 0.5 to 0.7 SD higher among families with more education, but with large variation within education groups. Changes in talk from 6 to 18 months highlighted that families from low educated backgrounds were decreasing the amount they spoke to their children (− 4219.54, 95% CI -6054.13, − 2384.95), compared to families from high educated backgrounds who remained relatively stable across this age period (− 369.13, 95% CI − 2344.57, 1606.30).

Conclusions: The socioeconomic word gap emerges between 12 and 18 months of age. Interventions to enhance maternal communication, child vocalisations and vocabulary development should begin prior to 18 months.

Keywords: Word gap, Parent talk, Inequality, Early childhood development, Language

Background

The emergence of socioeconomic inequalities in many areas of children's health and development is evident early in life [1, 2]. Understanding when and how these inequalities develop is a key question for researchers and policymakers because preventive interventions should be in place before health and development gaps become entrenched [3]. Hart and Risley famously coined the term '30 million word gap' by estimating through linear extrapolation of data collected from 10 to 36 months, that by age four, parents in the United States (US) who were on welfare had spoken 30 million words less to their child than parents with professional occupations [4]. In a 10-year follow up, they found these socioeconomic differences predicted subsequent verbal ability, receptive and expressive vocabulary, and academic achievement in grade 3 [5].

* Correspondence: mary.brushe@telethonkids.org.au
[1]Telethon Kids Institute, University of Western Australia, Level 15, 31 Flinders St, Adelaide, South Australia 5000, Australia
[2]School of Public Health, University of Adelaide, Level 9, Adelaide Health & Medical Science Building, 57 North Terrace, Adelaide, South Australia 5005, Australia
Full list of author information is available at the end of the article

The term '30 million word gap' has garnered enormous attention, with over 113 million google hits. In response, new technology has been developed [6] and considerable resources expended on initiatives across the world aiming to reduce the word gap. The Hart and Risley findings were based on a convenience sample of 42 families in Kansas, with only 6 families in the welfare category, compared to 13 families in the professional category and 23 families in the working class category. Furthermore, data were collected through researchers videotaping 1 h in the family's home per month, which may not be representative of the natural home environment. Language data were collected from 10 months of age onwards, limiting the understanding of critical language experiences during the first year of life. The validity and generalizability of Hart and Risley's findings have been widely debated [7–10].

Gilkerson and colleagues [11] attempted to overcome some of the limitations of Hart and Risley's work through the use of newly developed speech recognition technology, Language Environment Analysis (LENA). Researchers were able to objectively measure a family's home language environment to capture the number of words children heard over a day. The study involved 329 English speaking families with children aged between 2 and 48 months, from Denver. Families completed LENA recording days once a month for 6 months and a subset of 59 families completed monthly recording days for an additional 32 months. Their results estimated a 4-million-word gap by age 4 between mothers with some high school vs. those with a college degree.

The Language in Little Ones (LiLO) study is a prospective study of Australian families aiming to understand maternal education differences in the number of words children hear and speak in the home environment during the first 5 years of life. The LiLO study started collecting language data in the home, involving day-long recordings, when the children were 6 months old with data collection occurring every 6 months, until their first year of schooling, around age 5. We previously reported that when children were 6 and 12 months old there were no meaningful differences in any measure of parent-child talk between maternal education groups [12]. There was large variability, with high and low talkers within both education groups.

The present study includes new data from the LiLO study when children were 18 months old. This is an important age in children's language development when they are beginning to expand their vocabularies. Here we report all data currently available from the LiLO study including the number of adult words spoken to the child, number of child vocalizations and number of conversations between adult and child over a day when the children are 6, 12, and 18 months old by levels of maternal education.

Methods

Study design

The LiLO study commenced recording parent-child talk when children were 6 months old, with repeated measures every 6 months until child age 5 years. LiLO was explicitly designed to maximize contrasts across maternal education groups, by stratifying recruitment into a low educated group (mothers without any post-secondary school qualifications), and a high educated group (mothers with a bachelor's degree at minimum). At each wave of data collection, families undertook day-long (16-h) audio recordings. A $10 supermarket voucher was provided to families as compensation after each wave.

Participants

Recruitment occurred within Adelaide and Port Pirie in South Australia, Bunbury in Western Australia and the Gold Coast in Queensland between April 1, 2017 and July 31, 2019. Expecting mothers were approached at public hospitals while waiting for their antenatal appointments. Additionally, postnatal recruitment occurred at Child and Family Health Services during early parenting groups and drop-in clinics in Adelaide, Port Pirie and Bunbury. Mothers were also approached at local shopping centres, council-run immunization clinics, community playgroups, children's centres and libraries across all locations. Families were excluded if they did not speak English in the home or if the mother's level of education did not fall within the low or high educated categories. They were also excluded if their child was part of a multiple birth, was born premature (< 37 weeks), had a diagnosed cause of language impairment (e.g. hearing impairment, Down Syndrome, Cerebral Palsy) or was born outside the date range of January 1, 2017 and December 31, 2017.

Figure 1 provides a detailed flow chart of participant numbers across the first three waves. As is common in prospective studies, there was difficulty in attempting to recruit socioeconomically disadvantaged groups [13]. At the first wave, only 65 low educated families were participating in the study despite extensive and exhaustive recruitment efforts. To boost sample size among the low educated, we extended the recruitment timelines and locations which meant families could join the longitudinal study even if they had missed the first or second wave of data collection. An additional 35 low educated families joined the study and only seven families (4 low educated; 3 high educated) had withdrawn since the study started. A home visit occurred with each family within 2 months of the child's 6, 12 and 18 month birthdates. Data collection procedures have been previously described, with processes remaining consistent at each wave for all families in the study [12].
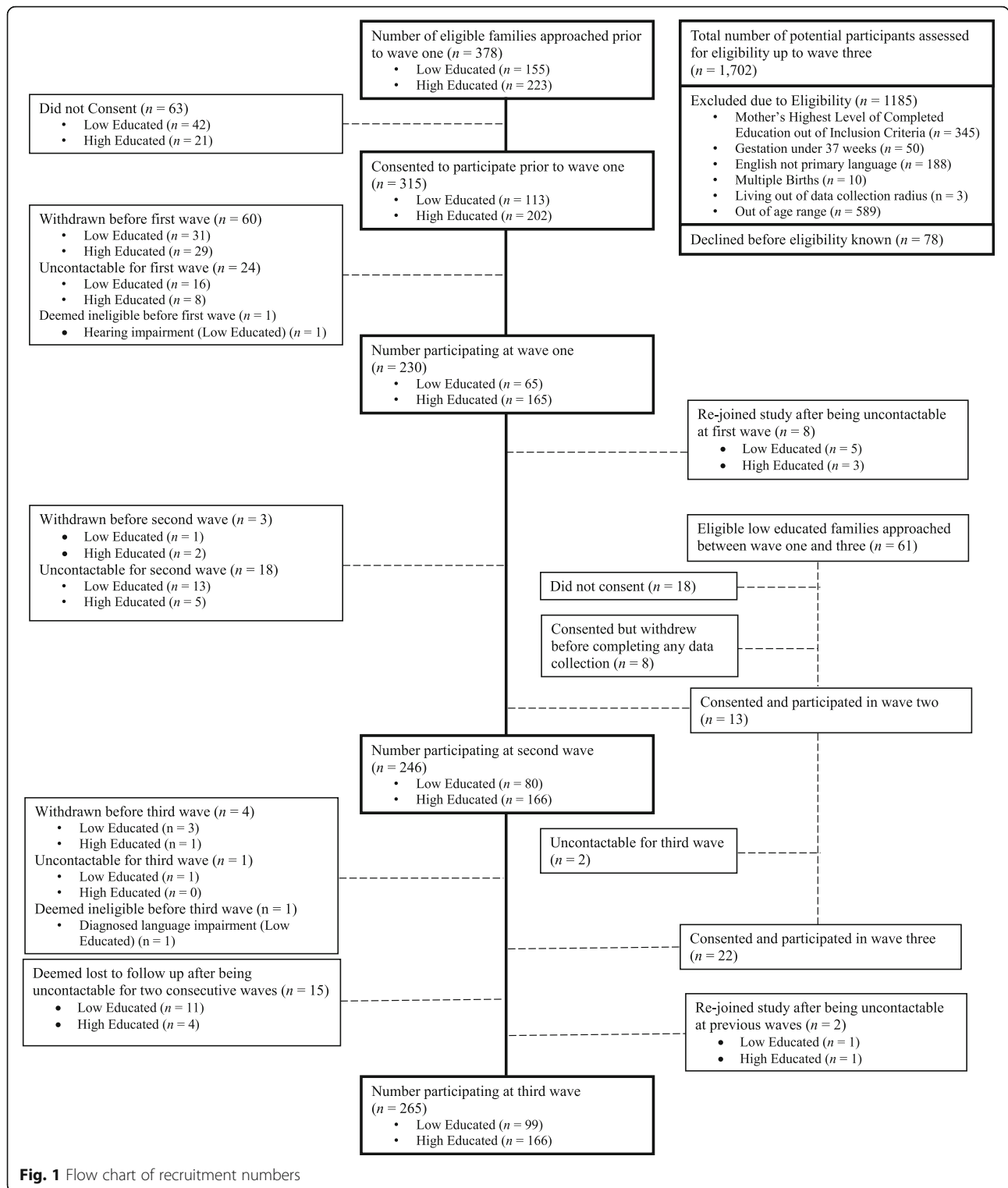
Number of eligible families approached prior to wave one (n = 378)
- Low Educated (n = 155)
- High Educated (n = 223)

Total number of potential participants assessed for eligibility up to wave three (n = 1,702)

Excluded due to Eligibility (n = 1185)
- Mother's Highest Level of Completed Education out of Inclusion Criteria (n = 345)
- Gestation under 37 weeks (n = 50)
- English not primary language (n = 188)
- Multiple Births (n = 10)
- Living out of data collection radius (n = 3)
- Out of age range (n = 589)

Declined before eligibility known (n = 78)

Did not Consent (n = 63)
- Low Educated (n = 42)
- High Educated (n = 21)

Consented to participate prior to wave one (n = 315)
- Low Educated (n = 113)
- High Educated (n = 202)

Withdrawn before first wave (n = 60)
- Low Educated (n = 31)
- High Educated (n = 29)
Uncontactable for first wave (n = 24)
- Low Educated (n = 16)
- High Educated (n = 8)
Deemed ineligible before first wave (n = 1)
- Hearing impairment (Low Educated) (n = 1)

Number participating at wave one (n = 230)
- Low Educated (n = 65)
- High Educated (n = 165)

Re-joined study after being uncontactable at first wave (n = 8)
- Low Educated (n = 5)
- High Educated (n = 3)

Withdrawn before second wave (n = 3)
- Low Educated (n = 1)
- High Educated (n = 2)
Uncontactable for second wave (n = 18)
- Low Educated (n = 13)
- High Educated (n = 5)

Eligible low educated families approached between wave one and three (n = 61)

Did not consent (n = 18)

Consented but withdrew before completing any data collection (n = 8)

Consented and participated in wave two (n = 13)

Number participating at second wave (n = 246)
- Low Educated (n = 80)
- High Educated (n = 166)

Withdrawn before third wave (n = 4)
- Low Educated (n = 3)
- High Educated (n = 1)
Uncontactable for third wave (n = 1)
- Low Educated (n = 1)
- High Educated (n = 0)
Deemed ineligible before third wave (n = 1)
- Diagnosed language impairment (Low Educated) (n = 1)

Uncontactable for third wave (n = 2)

Consented and participated in wave three (n = 22)

Deemed lost to follow up after being uncontactable for two consecutive waves (n = 15)
- Low Educated (n = 11)
- High Educated (n = 4)

Re-joined study after being uncontactable at previous waves (n = 2)
- Low Educated (n = 1)
- High Educated (n = 1)

Number participating at third wave (n = 265)
- Low Educated (n = 99)
- High Educated (n = 166)

Fig. 1 Flow chart of recruitment numbers

Measures

The Language Environment Analysis (LENA) system was used to capture the child's home language environment. The LENA technology comprises a digital language processor (DLP) and LENA computer software to automatically process the audio captured through the DLP using algorithmic analysis of the acoustic properties in the speech signal [6, 14, 15]. Three key LENA measures were used in this study: adult word count (AWC; the number of adult words

spoken to the child), child vocalization count (CVC; the number of speech-related sounds made by the child) and conversational turn count (CT; the number of alternations between adult and child occurring within at least 5 s of each other). Home activity diaries were also completed by the parents outlining the activities of the child, by the hour, throughout the recording day. Total word counts, from the LENA software, were used in the analysis when the full 16-h recording was completed or the activity diaries confirmed that the LENA device was turned off when the child went to sleep. Adjusted word counts were calculated if the LENA device was turned off prior to child's bedtime, whereby average hourly counts were added to the total reported word count to take the total recording time up until the child fell asleep, as reported by the parents in the home activity diaries. Adjusted word counts were only used for one low educated family in wave 1 and one high educated family in wave 3. Reliability testing by the LENA Foundation has reported high levels of agreement between human transcribers and LENA system classification [15].

Statistical approach

Parent-child talk variables were modelled using random effects longitudinal models using the *xtmixed* command in Stata, to understand changes in adult word counts, child vocalizations count, and conversational turn counts according to maternal education, from child ages 6, 12 and 18 months old. The interaction of mother's education and wave of data collection was included as the only predictor in the model to identify how changes over time differ between education groups. The parameters were computed using the expectation maximisation (EM) algorithm. To identify differences between maternal education groups and their word counts across each wave, we used the *margins* command in Stata to calculate the predicted means for low and high educated families at each time point and plotted their mean word counts and 95% confidence intervals across waves in Fig. 2, 3 and 4. A comparison of means from the observed data and the computed model is provided in the supplementary appendices (See Additional file 1). Effect sizes were also calculated using Cohen's *d* [16].

To ensure the addition of the extra 35 low educated families at waves 2 and 3 did not affect the results, we undertook a sensitivity analysis that only included families who had participated since wave 1. Details of the sensitivity analyses are also provided in the supplementary appendices (See Additional file 2). All analyses and figures were conducted using Stata version 16 [17].

Results

Data for the first three waves were collected between August 1, 2017 and July 31, 2019. Of those families actively participating in the study, LENA data was not available for 11 families across the three waves. This was due to nine families at wave three skipping their visit due to personal reasons and for two families the LENA device malfunctioned (one at wave 1 and one at wave 3). One family was deemed ineligible due to a diagnosed cause of language impairment at the third wave, so their data was retrospectively removed.

The sample varied slightly across waves due to the increase in participant numbers, as shown in Table 1. The final analysis sample consisted of 163 families in the high
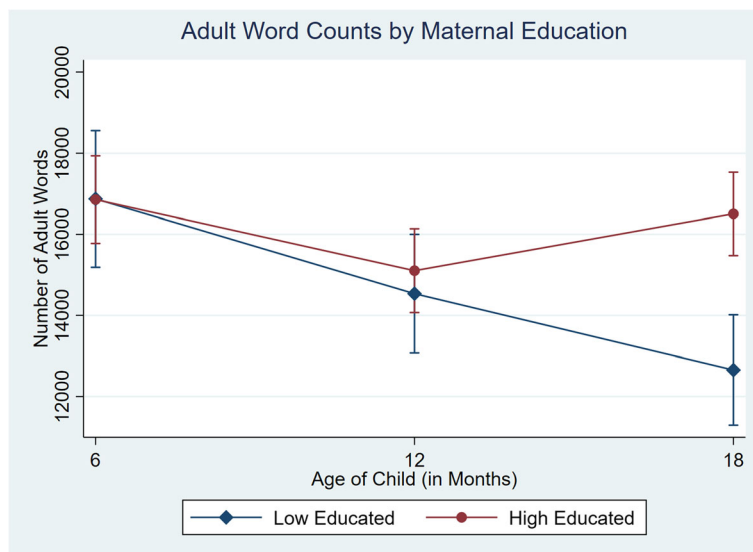

Fig. 2 Predicted mean adult word count and 95% CI by maternal education across 6 month, 12 month and 18 month wave of data collection

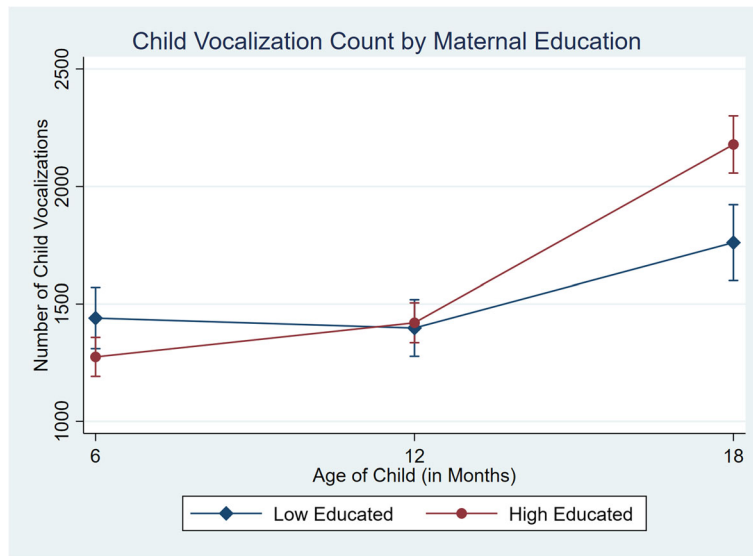Brushe *et al. BMC Pediatrics*   (2021) 21:247

Page 5 of 9

Fig. 3 Predicted mean child vocalizations count and 95% CI by maternal education across 6 month, 12 month and 18 month wave of data collection

educated group and 92 families in the low educated group. The average age of the mother at childbirth was 31.28 years and 85% of mother's were employed prior to their pregnancy. Just over half the children in the sample were first born and 54% were female.

Table 2 shows the results from the random effects model that estimates the interaction between mother's education and wave of data collection on the three LENA measures: adult word counts, child vocalization counts and conversational turn counts. The coefficient demonstrates the changes in growth for both low and high educated groups as compared to the 6 month baseline for the low educated group. As can be noted from the model, for adult word counts, families from low educated backgrounds were talking 4219.54 words less, 95% CI (− 6054.13, − 2384.95) to their children by 18 months, compared to high educated mothers who remained relatively stable across waves with only 369.13 fewer words 95% CI (− 2344.57, 1606.30) by 18 months. For child vocalization counts, the model demonstrates children
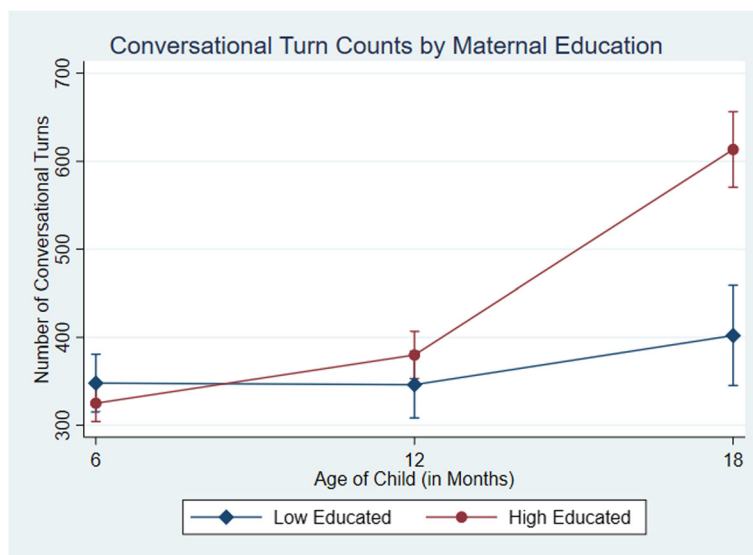

Fig. 4 Predicted mean conversational turns count and 95% CI by maternal education across 6 month, 12 month and 18 month wave of data collection

Brushe *et al. BMC Pediatrics*   (2021) 21:247

Page 6 of 9

Table 1 Sociodemographic Characteristics of the Sample

	6 month Data Collection (N = 228)	12 month Data Collection (N = 245)	18 month Data Collection (N = 255)
Child			
Age, mo, mean (SD)	5.82 (0.58)	11.99 (0.51)	18.02 (0.56)
Girls, n (%)	122 (53)	130 (53)	136 (53)
Gestation, wk., mean (SD)	39.1 (1.35)	39.18 (1.50)	39.19 (1.49)
Firstborn, n (%)	128 (56)	131 (53)	136 (53)
Mother			
Highest level of completed education, University, n (%)	164 (72)	166 (68)	163 (64)
Age at childbirth, y, mean (SD)	31.36 (4.42)	31.22 (4.57)	31.28 (4.84)
Working up until pregnancy, yes, n (%)	199 (87)	211 (86)	217 (85)

from both the low (320.74, 95% CI 126.61, 514.88) and high (739.04, 95% CI 560.66, 917.53) educated groups increased their number of vocalizations by 18 months, but high educated children grew their vocalizations at a faster rate. For conversational turn counts, both high and low educated families had little growth between the first and second wave, however between the ages of 12 and 18 months, growth in turns between adult and child for the high educated group (265.62, 95% CI 211.52, 319.73) exceeded that of the low educated group (54.22, 95% CI

-5.54, 113.98). The 95% confidence intervals in the models highlight large variability in growth across waves for both groups. However, on average low educated adults are talking less to their children by 18 months.

The graphs in Figs. 2, 3 and 4 depict the predicted mean and 95% confidence intervals for each measure of talk by maternal education groups at 6, 12 and 18 months of age. The figures show the emergence of the word gap for the number of adult words, child vocalizations and conversational turns by the time children were

Table 2 Random effects model estimates for LENA measures across maternal education groups

	Coef.	p	95% CI	
Adult Word Counts				
Number of adult words at 6 months among low educated = 16,872.86				
Low Educated at 12 months	-2336.90	0.016	-4243.32,	− 430.48
Low Educated at 18 months	−4219.54	0.000	−6054.13,	−2384.95
High Educated at 6 months	−16.64	0.987	−2019.84,	1986.55
High Educated at 12 months	− 1768.77	0.080	− 3746.90,	209.36
High Educated at 18 months	−369.13	0.714	−2344.57,	1606.30
Child Vocalisations Counts				
Number of child vocalizations at 6 months among low educated = 1440.28				
Low Educated at 12 months	−42.15	0.553	− 181.41,	97.11
Low Educated at 18 months	320.74	0.001	126.61,	514.88
High Educated at 6 months	− 165.36	0.036	−319.98,	−10.74
High Educated at 12 months	− 19.97	0.801	− 175.49,	135.54
High Educated at 18 months	739.04	0.000	560.55,	917.53
Conversational Turn Counts				
Number of conversational turns at 6 months among low educated = 347.90				
Low Educated at 12 months	−1.79	0.935	−44.76,	41.16
Low Educated at 18 months	54.22	0.075	−5.54,	113.98
High Educated at 6 months	−22.92	0.248	−61.82,	15.98
High Educated at 12 months	31.99	0.139	−10.42,	74.41
High Educated at 18 months	265.62	0.000	211.52,	319.73

18 months old, in line with the results of the random effects model. For adult words spoken (Fig. 2) we found a difference of 17 words at 6 months, 568 words at 12 months and 3851 words at 18 months, with families in the high educated group talking more at wave two and three. For child vocalizations (Fig. 3) children from the low educated group were vocalizing slightly more, with a difference between groups of 166 vocalizations at 6 months. By 12 months, there were only 22 more vocalizations on average from children of high educated mothers and by 18 months, children in the high educated group had on average 418 more vocalizations. For conversational turns (Fig. 4), there were similar differences at 6 and 12 months with 24 and 34 turns between adult and child respectively. Families in the low educated group engaged in slightly more conversational turns at 6 months but families in the high educated group had more conversations at 12 months. Similarly, as with the adult words and child vocalizations, by 18 months the difference in conversational turns had grown to 212 turns with more in the highly educated group. Effects for mothers with higher education ranged from 0.5 SD for child vocalizations up to 0.7 SD for conversational turns. Sensitivity analysis that included only families who were observed at each time point did not change the results (See Additional file 1).

Discussion

These results demonstrate that the word gap between high and low educated mothers emerges between 12 and 18 months. The differences between high and low educated mothers were seen for adult words, child vocalizations and conversational turns with effects ranging from 0.5 for word counts to 0.7 SD for conversational turns. As well as understanding the emergence of mean differences in all measures of talk by 18 months, it is important to note the large variability within education groups. There are high and low talkers across the socioeconomic spectrum even though on average more educated mothers engaged in more talk. These results are generally consistent with Gilkerson et al. [11], who reported more talk among high educated mothers in the aggregated age band from 20 to 26 months.

When considering the implementation of interventions to support the home language environments of infants and toddlers, these results suggest a proportionate universalist approach [18] may be more appropriate, whereby services are universally available but designed with a scale and intensity that is proportionate to the nature of disadvantage. While there is a mean difference between education groups at 18 months, there is also large variability in parent-child talk in both education groups, hence targeting interventions only towards low educated families would miss a large proportion of children who are experiencing lower levels of language stimulation in the home among better educated mothers. Targeting of interventions to particular subpopulations presents challenges in reducing inequalities in early childhood development [2].

A limitation of the current study is the differences in sample size across the education groups, with fewer families participating in the low educated group than originally planned. Nonetheless, the low educated group is 10 times larger than Hart and Risley's welfare group and twice as large as Gilkerson's et al. some high school group. Numerous strategies were employed to encourage participation. However fewer mothers were identified as eligible in this group resulting in a lower recruitment rate. At later waves an additional 35 families were recruited to the low education group despite missing early waves of data collection and recruitment will continue until our target sample size is reached.

The findings provide support for the existence of a socioeconomic word gap and that this gap emerges between 12 and 18 months of age. However, longer term data are required to quantify the size of the word gap by age 4. Key strengths of the LiLO study are that data collection began when children were 6 months old and it captures day-long audio recordings, compared to Hart and Risley who only captured 1-h of data in the early evening and did not begin data collection until 10-months old. Each family in the LiLO study is also followed longitudinally, unlike only the small subset of families from the Gilkerson et al. study. Additionally, the larger sample, compared to both Hart and Risley and Gilkerson et al., and the use of the LENA technology means LiLO is well placed to continue quantifying the socioeconomic disparities in talk during the first 5 years of life. Importantly, data were from a population-based sample purposely designed to maximise education exposure contrasts as has been recommended by leading methodologists [19]. These results are likely to be generalizable to the English-speaking Australian population, and probably other English-speaking populations, although there may be ethnic and cultural differences that were not examined in this study. Future LiLO research will consider whether trajectories of talk influence later developmental outcomes and how this differs for maternal education groups. It will also be important to monitor the large variation within the two education groups to see if it is maintained as children age, and if an environment of high talking among low educated families is associated with better child development outcomes.

Brushe *et al. BMC Pediatrics* (2021) 21:247

Page 8 of 9

Conclusion

These results from the LiLO study suggest a socioeconomic word gap emerges between the ages of 12 and 18 months. Families from low educated backgrounds decreased the amount they spoke to their children between 6 and 18 months, compared to families from high educated backgrounds whose quantity of talk remained relatively stable across the same period. This is the first study to have used an objective measure of a child's home language environment and been able to provide insight into the timing of the divergence of parent-child talk between maternal education groups. This finding suggests the implementation of proportionate universal programs that encourage parents to talk more to their child should occur prior to 18 months of age.

Abbreviations
LENA: Language ENvironment Analysis; LiLO: Language in Little Ones; DLP: Digital Language Processer; AWC: Adult Word Count; CVC: Child Vocalization Count; CT: Conversational Turn

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12887-021-02712-1.

Additional file 1: Supplementary Appendix. Observed Means vs Computed Means. To compare the observed means in the raw data and the predicted means computed using the *margins* command.

Additional file 2: Supplementary Appendix. Sensitivity Analysis. To ensure the addition of the extra 35 low educated families at waves 2 and 3 did not affect the results, we undertook a sensitivity analysis which only included families who had participated since wave 1.

Authors' contributions
SB, SR, JL, EM are chief investigators on the project and conceptualized the study. MB coordinated and supervised data collection for the study and wrote the first draft of the manuscript. JL, SB and MM conceptualized the analysis plan. MM and MB carried out the analysis. All authors contributed to the interpretation of findings and reviewed and revised the manuscript. Finally, all authors read and approved the final manuscript.

Availability of data and materials
The datasets generated and analysed during the current study are not publicly available due to lack of informed consent for data sharing at the time of collection, but are available from the corresponding author on reasonable request. For further information on the data and materials used in this study, please contact the corresponding author.

Declarations

Ethics approval and consent to participate
The Women's and Children's Health Network (HREC/16/WCHN/190) and the University of Western Australia's (RA/4/1/8825) Human Research Ethics Committee approved this study, and written informed consent was provided by all participants and by the parent or guardian for participants under 16 years old. All methods were carried out in accordance with relevant guidelines and regulations.

Consent for publication
Not applicable.

Competing interests
The authors declare they have no competing interests.

Author details
[1]Telethon Kids Institute, University of Western Australia, Level 15, 31 Flinders St, Adelaide, South Australia 5000, Australia. [2]School of Public Health, University of Adelaide, Level 9, Adelaide Health & Medical Science Building, 57 North Terrace, Adelaide, South Australia 5005, Australia. [3]Population Health Sciences, Bristol Medical School, University of Bristol, First Floor, 5 Tyndall Avenue, Bristol B28 1UD, UK. [4]Menzies Health Institute Queensland G40 Griffith Health Centre, Griffith University, Level 8.86 Gold Coast Campus, Southport, Queensland 4222, Australia. [5]Department of Education, University of Oxford, 15 Norham Gardens, Oxford OX2 6PY, UK.

References
1. Brinkman S, Gialamas A, Rahman A, Mittinty MN, Gregory TA, Silburn S, et al. Jurisdictional, socioeconomic and gender inequalities in child health and development: analysis of a national census of 5-year-olds in Australia. BMJ Open. 2012;2(5):e001075. https://doi.org/10.1136/bmjopen-2012-001075.
2. Lynch JW, Law C, Brinkman S, Chittleborough C, Sawyer M. Inequalities in child healthy development: some challenges for effective implementation. Soc Sci Med. 2010;71(7):1219–374. https://doi.org/10.1016/j.socscimed.2010.07.008.
3. Lynch J. Conference keynote. What can early interventions really achieve, and how will we know? Family matters: newsletter of the Australian Institute of Family. Studies. 2017;99:1.
4. Hart B, Risley TR. Meaningful differences in the everyday experience of young American children. Baltimore: Paul H. Brookes Publishing Co.; 1995.
5. Walker D, Greenwood C, Hart B, Carta J. Prediction of school outcomes based on early language production and socioeconomic factors. Child Dev. 1994;65(2):606–21. https://doi.org/10.2307/1131404.
6. Gilkerson J, Richards JA. The Power of Talk: Impact of Adult Talk, Conversational Turns, and TV During the Critical 0–4 Years of Child Development. Boulder: LENA Research Foundation; 2009. Contract No.: LENA Technical Report LTR-01-2
7. Golinkoff RM, Hoff E, Rowe ML, Tamis-LeMonda CS, Hirsh-Pasek K. Language matters: denying the existence of the 30-million-word gap has serious consequences. Child Dev. 2019;90(3):985–92. https://doi.org/10.1111/cdev.13128.
8. Sperry DE, Sperry LL, Miller PJ. Language Does Matter: But There is More to Language Than Vocabulary and Directed Speech. Child Dev. 2018;0:0. https://doi.org/10.1111/cdev.13125.
9. Baugh J. Meaning-less differences: exposing fallacies and flaws in "the word gap" hypothesis that conceal a dangerous "language trap" for low-income American families and their children. Int Multi Res J. 2017;11(1):39–51. https://doi.org/10.1080/19313152.2016.1258189.
10. Purpura DJ. Language clearly matters; Methods Matter Too. Child Dev. 2019; 90(6):1839–46. https://doi.org/10.1111/cdev.13327.
11. Gilkerson J, Richards JA, Warren SF, Montgomery JK, Greenwood CR, Oller DK, et al. Mapping the early language environment using all-day recordings and automated analysis. Am J Speech-Language Pathol. 2017;26(2):248–65. https://doi.org/10.1044/2016_AJSLP-15-0169.
12. Brushe ME, Lynch JW, Reilly S, Melhuish E, Brinkman SA. How many words are Australian children hearing in the first year of life? BMC Pediatr. 2020; 20(1):52. https://doi.org/10.1186/s12887-020-1946-0.

13. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol. 2007;17:643–53.

14. Xu D, Richards JA, Gilkerson J. Automated analysis of child phonetic production using naturalistic recordings. J Speech Language Hearing Res. 2014;57(5):1638–50. https://doi.org/10.1044/2014_jslhr-s-13-0037.

15. Xu D, Yapanel U, Gray S. Reliability of the LENA language environment analysis system in young Children's natural home environment. Boulder: LENA Foundation; 2009. 16 p.

16. Cohen J. Statistical power analyses for the Behavioural sciences. 2nd ed. New Jersey: Lawrence Erlbaum Associates; 1988.

17. StataCorp. Stata Statistical Software: Release 16. TX: StataCorp LLC: College Station 2019.

18. Marmot M. Fair society, healthy lives: the marmot review; 2010. Available from: http://www.marmotreview.org/AssetLibrary/pdfs/Reports/FairSocietyHealthyLives.pdf.

19. Rothman KJ, Gallacher JE, Hatch EE. Why representativeness should be avoided. Int J Epidemiol. 2013;42(4):1012–4. https://doi.org/10.1093/ije/dys223.

Publisher's Note