ASHA American Speech-Language-Hearing Association

JSLHR

Review Article

What the Evidence Does (and Does Not) Show for the Centers for Disease Control and Prevention Child Development Milestones: An Illustrative Example Using Expressive Vocabulary

Megan Y. Roberts,[a,b] Bailey J. Sone,[a] Maranda K. Jones,[a] Murielle Standley,[a] Tracy Conner,[a] E. Debbie Lee,[a] Elizabeth S. Norton,[a,b] Judith Roman,[a] Marisha Speights,[a] Rylie Young,[a] and Adriana Weisleder[a]

[a]Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL [b]Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL

ARTICLE INFO

Article History:
Received January 10, 2023
Revision received March 28, 2023
Accepted May 22, 2023

Editor-in-Chief: Julie A. Washington
Editor: Mabel L. Rice

https://doi.org/10.1044/2023_JSLHR-23-00020

ABSTRACT

Purpose: Child development milestones are a critical tool for pediatricians and caregivers to use for developmental surveillance. Following review and selection by a panel of subject matter experts, the Centers for Disease Control and Prevention (CDC) published a revised list of milestones across multiple domains of development. Using expressive vocabulary, a key indicator of language development, as an illustrative example, the purpose of this brief review is to evaluate the evidence used to establish the CDC developmental milestones and determine whether the samples used to establish these milestones are representative of U.S. children.

Method: Authors reviewed the methods and evidence cited to determine the CDC milestones. First, authors identified each language/communication milestone that measured expressive vocabulary as number of words, followed by review of the sources cited in support of each extracted milestone. Then, data related to both milestones and sample characteristics were extracted and compiled as well as compared with data from a validated parent report measure of expressive vocabulary, the MacArthur–Bates Communication Development Inventories.

Results: Results indicated that evidence was conflicting, misaligned, or missing for the selected CDC expressive vocabulary milestones. This review also indicated that the samples used to determine the selected CDC expressive vocabulary milestones are not representative of U.S. children.

Conclusion: The striking paucity of evidence supporting the new CDC milestones for expressive vocabulary illustrates the critical need for future research in this area to establish more accurate milestones for U.S. children, with a focus on culturally inclusive large-scale data.

Child development milestones are a key tool for developmental surveillance used by pediatricians and caregivers, as the American Academy of Pediatrics (Lipkin et al., 2020) recommends using milestones to ensure timely screening, diagnosis, and intervention for developmental delays or disabilities. The Centers for Disease Control and Prevention (CDC) published lists of developmental milestones for children for multiple domains of development: social/emotional, language/communication, cognitive, and motor (CDC, 2022; Zubler et al., 2022). Accurate milestones are critical to aid pediatricians and caregivers in the process of developmental screening and help ensure identification of atypical developmental patterns, such that

Correspondence to Megan Y. Roberts: megan.y.roberts@northwestern.edu. ***Disclosure:*** *Adriana Weisleder is an unpaid member of the MCDI Advisory Board. The other authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

intervention begins as early as possible. In 2022, the CDC revised these milestones with a review of the empirical evidence (CDC, 2022; Zubler et al., 2022). Although milestones checklists are not meant to replace validated screening or diagnostic tools (CDC, 2022), they are used by pediatricians during health supervision visits to identify developmental concerns. As such, these milestones should (a) be supported by strong empirical evidence, (b) be representative of the population for which they are intended (in this case, U.S. children), and (c) offer clear cutoffs to aid in clinical decision making. In this article, we focus on the revised language milestones from the CDC (2022) and the lack of alignment with published, peer-reviewed research.

To determine the developmental milestones as part of the CDC milestone revision, a panel of subject matter experts was convened, led by Zubler et al. (2022). The panel included pediatricians with general, developmental-behavioral, and neurodevelopmental expertise; child and developmental psychologists; and a professor of special education and early intervention, yet lacked representation from speech-language pathologists. This panel conducted a literature review in conjunction with a review of developmental resources (i.e., screening and diagnostic tools) by developmental experts. They reported that 80% of their revised milestones were supported by normative data from one or more sources, an improvement from the CDC milestones published in 2004 (Zubler et al., 2022). Another notable improvement since 2004 is the published article detailing the process of establishing the developmental milestones, an important step toward increasing transparency between the public and scientists. Indeed, publication of the new milestones generated vigorous public discussion, and some professional organizations, including those that represent language and communication, highlighted the underrepresentation of speech-language pathologists during the process of establishing the developmental milestones and voiced disagreement with the accuracy of the final milestones (American Speech-Language-Hearing Association [ASHA], 2022; The Informed SLP Team, 2022). The current investigation, conducted by an independent group of speech-language pathology experts at Northwestern University, came about in response to this discussion. In particular, we sought to provide an independent investigation of the evidence supporting the selection of language and communication milestones.

Not only should normative data exist to support the selection of milestones, but these normative data should be representative of the population for which the milestones are intended to provide developmental surveillance. This is especially important for milestone checklists published by government agencies such as the CDC, since they are meant as a guide for the entire U.S. population. To support the milestones, Zubler et al. (2022) included normative data garnered from studies both in and outside the United States, including some non–English-speaking countries. They did not provide information about how well the samples in the cited sources represent the diversity of U.S. children with regards to race, ethnicity, and socioeconomic status. In addition, the article did not mention whether the normative data supporting the milestones included dual language learners, who may have different trajectories of language development (Bialystok et al., 2009; De Houwer, 2021; Genesee, 2008).

Furthermore, milestones should use clear and consistent cutoffs to aid in clinical decision making. The 2004 CDC milestones represented the age at which 50% of children demonstrate the skill/behavior in question, which would be equivalent to using a cutoff at the 50th percentile (Zubler et al., 2022). A primary change in the 2022 CDC milestones is that the milestones represent when 75% or more children are expected to achieve the skill/behavior at a given age (i.e., equivalent to using a cutoff of the 25th percentile or below; Zubler et al., 2022). This criterion is intended to reduce pediatricians taking a "wait and see" approach, which is currently common, given that falling behind 75% of children on any given milestone would be more indicative of possible developmental delay or disorder than falling slightly below average (Zubler et al., 2022). For the CDC milestones to effectively meet this criterion, it is critical that the existing normative data support that expected age of acquisition for each milestone (i.e., that it aligns with the 25th percentile).

Purpose and Guiding Questions

The purpose of this article is to evaluate the evidence used to establish the CDC developmental milestones for language and communication, using normative data for children's expressive vocabulary as an illustrative example. We chose to focus on the items measuring expressive vocabulary as number of words because (a) they are a key indicator of expressive language development (Dale & Patterson, 2017; Manning et al., 2019; Paul, 1996; Rescorla, 1989), (b) these are some of the most frequently noticed and used milestones for evaluating children's language and communication development (Dale & Patterson, 2017; Rescorla, 1989), and (c) they have been extensively researched within the fields of early language acquisition, pediatrics, and communication sciences and disorders (Fischel et al., 1989; Fisher, 2017; Frank et al., 2021; Gilkerson et al., 2018; Hammer et al., 2017). We address the following guiding questions (GQs):

GQ1: Do the sources cited for the selected CDC expressive vocabulary milestones provide any evidence that supports the expected age of acquisition for these milestones?

GQ2: What are the participant characteristics for the samples from which the selected CDC expressive vocabulary milestones were derived, and are they representative of the population of U.S. children?

GQ3: Do other sources of normative data on children's vocabulary development provide evidence consistent with the selected CDC expressive vocabulary milestones?

Method

Identification of Evidence

Authors reviewed the methods and evidence cited in Zubler et al.'s (2022) investigation. The first step was the identification and extraction of each language or communication milestone that included number of words. We extracted three milestones related to number of words. These CDC milestones include (a) tries to say one or two words besides mama or dada such as "ba" for ball or "da" for dog at 15 months; (b) tries to say ≥ 3 words besides mama or dada at 18 months; and (c) says ∼50 words at 30 months (Zubler et al., 2022). For the first milestone, the number of words acquired may range from one to four words, and for the second milestone, the number of words acquired may range from three to five words (depending on whether the child does or does not say "mama" or "dada"). In order to distinguish between the first and second milestones, we operationalized the 15-month milestone to include any data relevant to the acquisition of 1–3 words and the 18-month milestone to include any data relevant to the acquisition of 4+ words. The sources cited for each of these milestones were reviewed by the second author, and data related to all milestones and sample characteristics were extracted and compiled. Then, we evaluated whether the extracted data supported the CDC's decision to establish a milestone at the corresponding age (i.e., 15, 18, or 30 months) using the 25th percentile as a threshold. The sources cited were rated as supporting the CDC's decision if the source (a) reported data specific to the construct of the milestone, (b) reported data specific to the corresponding age or 25th percentile, and (c) the 25th percentile aligned with the corresponding milestone age (rounded to the nearest whole number). Given that the purpose of this article was to verify the sources used by Zubler et al. (2022) to establish the new milestones, we did not conduct an additional systematic search. Additionally, Zubler et al. (2022) increased transparency by using clear reporting of their methods such that it was not necessary to conduct a second search. However, we chose to compare data from these sources with data from the most commonly used and validated parent report instrument of expressive vocabulary, the MacArthur–Bates Communication Development Inventories (henceforth referred to as CDI; Fenson et al., 2007).

Evidence Verification

The third author reviewed all extracted data for verification. Discrepancies were discussed between this author and the author who completed the primary data extraction (second author). Consensus was reached on all data such that the reviewing authors were in 100% agreement.

Results

Guiding Question 1: Do the Sources Cited for the Selected CDC Expressive Vocabulary Milestones Provide Any Evidence That Supports the Expected Age of Acquisition for These Milestones?

Relevant findings are shown in Tables 1–3. Six sources were cited in total across the three CDC expressive vocabulary milestones. Although not all six sources were cited for all three milestones, we reviewed all sources for completeness.

CDC expressive vocabulary Milestone 1. None of the sources supported the CDC's decision to establish a milestone related to the production of 1–2 words besides mama and dada at 15 months using the 25th percentile as a threshold. Of the six sources, four of the sources reported data related to the construct of the milestone. Only one source included data for the 25th percentile (Gladstone et al., 2010), and that source suggests that this milestone is met at 18 months, not 15 months. All other data were reported using a threshold of the 50th percentile and suggest that this milestone may be met between 7 and 14 months (Ertem et al., 2018; Lancaster et al., 2018; Tamis-LeMonda et al., 1998).

CDC expressive vocabulary Milestone 2. Only one source supported the CDC's decision to establish a milestone related to the production of ≥ 3 words besides mama or dada at 18 months using the 25th percentile as a threshold (Sheldrick & Perrin, 2013). Of the six sources, four of the sources reported data related to the construct of the milestone. The one source that supported the decision to establish this milestone presented data for the 25th percentile (Sheldrick & Perrin, 2013). All other studies reported data for the 50th percentile and suggest that, at this threshold, the milestone may be reached between 13 and 16 months (Ertem et al., 2018; Lancaster et al., 2018).

Table 1. Analysis of references for the 15-month milestone in Zubler et al. (2022).

Milestone: tries to say one or two words besides mama or dada, like "ba" for ball or "da" for dog				
	Data reported			Does the data support the decision to establish this milestone at 15 months using a threshold of the 25th percentile?
Source	Item	Statistics	Location	
Accardo & Capute (2005)[a]				**NO** – Data are not specific enough to evaluate the construct of this milestone. Data are only presented for a total score on a developmental assessment.
Ertem et al. (2018)	Uses two meaningful words	12.5 months = 50th percentile	Page e284, Item 11	**NO** – Data are not reported at 15 months or the 25th percentile.
Gladstone et al. (2010)	Says two or more words besides mama/dada	14 months = 50th percentile 18 months = 25th percentile	Page e1000273, Figure 6	**NO** – Using a threshold of the 25th percentile, data suggest this milestone is reached at 18 months.
Lancaster et al. (2018)	Says two words, but words other than those used for mother and father Uses one word approximations Uses one definite word	14 months = 50th percentile 7.1–12.2 months = 50th percentile 10.7 months = 50th percentile	Supplemental Materials, Pages 120–121 and 126–127	**NO** – Data are not reported at 15 months or the 25th percentile.
Sheldrick & Perrin (2013)				**NO** – Data are not reported for this milestone.
Tamis-LeMonda et al. (1998)	First words in production	12.8 months = mean age	Page 687, Table 1	**NO** – Data are not reported at 15 months or the 25th percentile.

[a]Zubler et al. (2022) cites the assessment manual data were derived from the article reporting its standardization (Visintainer et al. 2004); Zubler et al. (2022) does not cite all six sources for each milestone, but we examined all six sources for all three milestones regardless if it was cited for a particular milestone.

CDC expressive vocabulary Milestone 3. Finally, none of the sources provided data to support that the production of 50 words is reached as late as 30 months. Only one of the six sources cited reported data related to the construct of the milestone, yet no data were reported for the 25th percentile. Tamis-Lamonda et al. (1998) suggest that the mean age of acquisition for 50 words is 17.9 months.

Guiding Question 2: What Are the Participant Characteristics for the Samples From Which the Selected CDC Expressive Vocabulary Milestones Were Derived, and Are They Representative of the Population of U.S. Children?

Relevant findings are shown in Table 4. All six sources provided some descriptive data for the sample of included participants. In general, the references cited by Zubler et al. (2022) for the selected CDC expressive vocabulary milestones came from three United States (Accardo & Capute, 2005; Sheldrick & Perrin, 2013; Tamis-LeMonda et al., 1998) and three international samples (Ertem et al., 2018; Gladstone et al., 2010; Lancaster et al., 2018). Two of the studies from the United States reported race (Accardo & Capute, 2005; Sheldrick & Perrin, 2013), and one of these two studies included a sample that was predominantly White (Sheldrick & Perrin, 2013). Only one study from the United States reported ethnicity (Sheldrick & Perrin, 2013). All of the studies from the United States reported data on socioeconomic status (maternal education, family income, or a composite), and all had overrepresentation of families from higher socioeconomic strata. All studies from the United States included samples of participants who spoke only English (Accardo & Capute, 2005; Sheldrick & Perrin, 2013; Tamis-LeMonda et al., 1998). The data from outside the United States included languages other than English, but the participants' language backgrounds were not always reported. Critically, the single source that reports the 25th percentile for the production of ≥ 3 words besides mama or dada at 18 months (CDC expressive vocabulary Milestone 2) is from a U.S.-based sample of predominantly White, English-speaking children (Sheldrick & Perrin, 2013).

Table 2. Analysis of references for the 18-month milestone in Zubler et al. (2022).

Milestone: tries to say ≥ 3 words besides mama or dada				
	Data reported			Does the data support the decision to establish this milestone at 18 months using a threshold of the 25th percentile?
Source	Item	Statistics	Location in cited article	
Accardo & Capute (2005)[a]				**NO** – Data are not specific enough to evaluate the construct of this milestone. Data are only presented for a total score on a developmental assessment.
Ertem et al. (2018)	Uses four meaningful words	15.3 months = 50th percentile	Page e284, Item 13	**NO** – Data are not reported at 18 months or the 25th percentile.
Gladstone et al. (2010)				**NO** – Data are not reported for this milestone.
Lancaster et al. (2018)	Says more than three words Says 4+ words	13.5 months = 50th percentile 13.3–15.2 months = 50th percentile	Supplemental Materials, Pages 118–119 and 124–125	**NO** – Data are not reported at 18 months or the 25th percentile.
Sheldrick & Perrin (2013)	Names at least 5 familiar objects – like ball or milk	16.1 months = 50th percentile 17.5 months = 25th percentile	Page 583, Item 28	**YES** – Using a threshold of the 25th percentile, data suggest this milestone is reached at 17.5 months.
Tamis-LeMonda et al. (1998)				**NO** – Data are not reported for this milestone.

[a]Zubler et al. (2022) cites the assessment manual data were derived from the article reporting its standardization (Visintainer et al. 2004); Zubler et al. (2022) does not cite all six sources for each milestone, but we examined all six sources for all three milestones regardless if it was cited for a particular milestone.

Guiding Question 3: Do Other Sources of Normative Data on Children's Vocabulary Development Provide Evidence for the Selected CDC Expressive Vocabulary Milestones?

Relevant findings are shown in Table 5. For our third question, we chose to compare the CDC milestones with data from the most commonly used and validated parent-report instrument of expressive vocabulary, the CDI (Fenson et al., 2007). The CDI reports normative data on expressive vocabulary for U.S. English-speaking children between 9 and 30 months of age and provides percentile ranks for the number of words produced at each 1-month age band (see Fenson et al., 2007, as well as http://wordbank.stanford.edu/). Note that this source was not included in the original selection of the CDC expressive vocabulary milestones (Zubler et al., 2022). Therefore, in addition, we included data from the Ages and Stages Questionnaire–Third Edition (ASQ-3), which was included in the original selection of the CDC expressive vocabulary milestones (Squires & Bricker, 2009).

The norming sample for the Words and Gestures CDI, which covers the 8- to 18-month age range, included 1,089 U.S. children (50% female) who were identified as 72.9% White, 11.4% Black, 4.2% Asian, 5.7% Hispanic, and 5.7% mixed or other race/ethnicity (Words and Gestures CDI; Fenson et al., 2007). Maternal education was used as a proxy for socioeconomic status with 6.8% reporting some high school or less, 23.6% reporting a high school diploma, 25.1% some college education, and 44.5% a college diploma. The norming sample for the Words and Sentences CDI, which covers the 16- to 30-month age range, included 1,461 U.S. children (50% male) who were identified as 73.7% White, 9.7% Black, 2.7% Asian, 7.1% Hispanic, and 6.9% mixed or other race/ethnicity (Words and Sentences CDI; Fenson et al., 2007). Maternal education was used as a proxy for socioeconomic status with 8.2% reporting some high school or less, 24.1% reporting a high school diploma, 24.5% some college education, and 43.2% a college diploma. The CDI norming sample consists of a sample collected in 1992/93 (Fenson et al., 1993) and an updated sample collected and added to the original sample in 2007 (Fenson et al., 2007). The first normative sample included children for whom English was the primary language. Of that sample, 12.2% reported their children were exposed to a second language with "less frequency than English" (Fenson et al., 2007); the most frequent second language was Spanish (45.6%), with the other 54.4% divided among 29 different languages. In the updated sample, children were excluded if they were exposed to a language other than English for more than 12 hr per week (Fenson et al., 2007).

When we compared the number of words recommended by the CDC milestones (one or two words besides mama or dada at 15 months, three or more words besides mama or dada at 18 months, 50 words at 30 months) to the CDI normative data, the percentile rank for each of

Table 3. Analysis of references for the 30-month milestone in Zubler et al. (2022).

Milestone: says ~ 50 words				
	Data reported			Does the data support the decision to establish this milestone at 30 months using a threshold of the 25th percentile?
Source	Item	Statistics	Location in cited article	
Accardo & Capute (2005)[a]				**NO** – Data are not specific enough to evaluate the construct of this milestone. Data are only presented for a total score on a developmental assessment.
Ertem et al. (2018)				**NO** – Data are not reported for this milestone.
Gladstone et al. (2010)				**NO** – Data are not reported for this milestone.
Lancaster et al. (2018)				**NO** – Data are not reported for this milestone.
Sheldrick & Perrin (2013)				**NO** – Data are not reported for this milestone.
Tamis-LeMonda et al. (1998)	50 words in production	17.9 months = mean age	Page 687, Table 1	**NO** – Data are not reported at 30 months or the 25th percentile.

[a]Zubler et al. (2022) cites the assessment manual data were derived from the article reporting its standardization (Visintainer et al. 2004); Zubler et al. (2022) does not cite all six sources for each milestone, but we examined all six sources for all three milestones regardless if it was cited for a particular milestone.

these milestones was at the 15th percentile or lower. In other words, while the goal of the CDC milestones is to indicate when 75% of children have acquired a particular skill, data from the CDI suggests that none of the selected CDC expressive vocabulary milestones meet that goal. Rather, these milestones indicate when 85% or more of children have acquired that skill. In fact, when using the CDI, the 25th percentile for the 15-, 18-, and 30-month milestones for expressive vocabulary is considerably different than the CDC milestones. Specifically, a child would be in the 25th percentile on the CDI if they (a) produced one to two words besides mama and dada between 13 and 14 months (i.e., one to three words; 1–2 months earlier than the CDC milestone), (b) produced three words besides mama or dada at 15 months (3 months earlier than the CDC milestone), and (c) produced 50+ words at 18–20 months (10–12 months earlier than the CDC milestone).

Discussion

In this article, our goal was to provide an in-depth evaluation of the cited evidence for the CDC expressive vocabulary milestones that included consideration of participant characteristics and comparison to a commonly used measure of child expressive vocabulary. The published article by Zubler et al. (2022) is an important step toward increasing transparency with respect to the establishment of developmental milestones. However, our review of the evidence suggests that the existing normative data do not align with the selected CDC-expressive vocabulary milestones. First, the evidence cited for each milestone by Zubler et al. (2022) is scant. Only one milestone (18 months milestone: tries to say ≥ 3 words besides mama or dada) had any normative data as supporting evidence, and this supporting data came from only one of the six cited sources. Second, the sources are not representative of the U.S. population. Some sources are from studies conducted outside of the United States, with no consideration of how linguistic, cultural, and socioeconomic differences between the United States and other countries could affect the age at which these milestones are reached. On the other hand, samples from studies in the United States were not representative of the racial, ethnic, linguistic, and socioeconomic diversity of the U.S. population. Taken together, there is not enough evidence to support the new CDC-expressive vocabulary milestones for U.S. children when considering both the findings from the sources used to support the milestones and the participant characteristics included in these sources. Third, when compared with the CDI norming data, the largest database on caregiver-reported expressive vocabulary for children in the United States, none of the CDC milestones align with the 25th percentile. Instead, these milestones align with the 15th percentile or lower on the CDI. As such, the CDC milestones will fail to identify a large percentage of children who score between the < 15th and 25th percentile who would benefit from earlier identification by their pediatrician and a referral for further assessment. In summary, the CDC-expressive vocabulary milestones are not an optimal tool for developmental surveillance given the existing evidence.

Initiatives to improve developmental surveillance processes that relate specifically to communication must

Table 4. Participant characteristics from references cited in Zubler et al. (2022).

Source	*N*	Country	Male	Race[b]	Ethnicity[b]	Language	SES[b]
Accardo & Capute (2005)[a]	1,239	USA	53.2%	56.7% White 32.3% Black 11% Other	—	English	Maternal education: 3.7% less than high school 12.2% some high school 23.6% high school degree 44.4% college degree 15.7% postbaccalaureate 00.4% missing information
Ertem et al. (2018)	4,949	Argentina, India, South Africa, Turkey	59%	—	—	Spanish, Turkish, isiZulu, sePedi, seTswana, English, Marathi, Gujrathi, Hindi	Maternal education: 49% < 12 years 51% ≥ 12 years
Gladstone et al. (2010)	1,446	Malawi	—	—	—	—	Wealth quintile: 21% lowest 18% second 20% middle 21% fourth 20% highest
Lancaster et al. (2018)	21,083	Costa Rica, Nicaragua, Paraguay, Peru, Bangladesh, India, Indonesia, Kenya, Malawi, Tanzania	—	—	—	—	Compiled multiple data sets, SES not consistently specified for the whole group. Noted that sample came from "10 low-middle income countries (LMIC)."
Sheldrick & Perrin (2013)	469	USA	53%	73% White 10% Black 1% Native American 9% Asian 3% Other/multiple 4% Not indicated	14% Hispanic	English	Family income: 17% < $20,000 14% $20,000–$49,000 28% $50,000–$99,000 38% ≥ $100,000 4% not indicated
Tamis-LeMonda et al. (1998)	40	USA	43%	—	—	English	"Children came from relatively homogeneous, middle- to upper-middle-class intact households" *M* = 58.7, *SD* = 6.3, on the Hollingshead Four Factor Index of Social Status, 1975

Note. The em dashes signify sections of data that were not reported in the reviewed article. SES = socioeconomic status.

[a]Zubler et al. (2022) cites the assessment manual data were derived from the article reporting its standardization (Visintainer et al., 2004). [b]For comparison, we present data from the 2020 American Community Survey Census Data 5-year Estimates. Of the total U.S. population under the age of 5 years in 2020, 63.1% identified as White, 13.8% as Black or African American, 0.9% as American Indian or Alaska Native, 5.2% as Asian, 0.2% as Native Hawaiian and Other Pacific Islander, 6.3% as some other race, and 10.5% as more than one race. In addition, 48.6% identified as White alone (not Hispanic or Latino) and 25.6% identified as Hispanic or Latino. Of women who had a birth in the past 12 months, 11.01% had less than high school degree, 22.91% had a high school degree, 30.66% had some college or an associate's degree, 21.91% had a bachelor's degree, and 13.51% had a graduate or professional degree (American Community Survey, 2021).

Table 5. CDC milestones and CDI data.

Milestone	Source	Data reported			Does the data support the decision to establish this milestone at the corresponding age using a threshold of the 25th percentile?
		Item	Statistics	Location	
15 months: tries to say one or two words besides mama or dada, like "ba" for ball or "da" for dog	CDI	1–3 words produced	10–12 months = 50th percentile 13–14 months = 25th percentile 15 months = 10th–15th percentile	CDI WG Scoring manual, Page 119	**NO –** Using a threshold of the 25th percentile, data suggest this milestone is reached between 13 and 14 months. Children that do not reach this milestone until 15 months would be in the 10th–15th percentile.
	ASQ	Does your baby say three words, such as "Mama," "Dada" and "Baba"?	14 months = age expected[a]	14 month ASQ checklist	**NO –** The age expected for children to reach this milestone is 14 months.
18 months: tries to say ≥ 3 words besides mama or dada	CDI	4+ words produced	13 months = 50th percentile 15 months = 25th percentile 18 months = 5th–10th percentile	CDI WG Scoring manual, page 119	**NO** – Using a threshold of the 25th percentile, data suggest this milestone is reached at 15 months. Children that do not reach this milestone until 18 months would be in the 5th–10th percentile.
	ASQ	Does your baby say four or more words in addition to "Mama" and "Dada"?	14 months = age expected[a]	14-month ASQ checklist	**NO** – The age expected for children to reach this milestone is 14 months.
30 months: says 50 words	CDI – Words and Gestures	50 words produced	17 months = 50th percentile 18 months = 25th percentile	CDI WG Scoring manual, page 119	**NO** – Using a threshold of the 25th percentile, data suggest this milestone is reached at 18 months.
	CDI – Words and Sentences	50 words produced	16 months = 50th percentile 20 months = 25th percentile 30 months ≤ 5th percentile	CDI WS Scoring manual, page 125	**NO** – Using a threshold of the 25th percentile, data suggest this milestone is reached at 20 months. Children that do not reach this milestone until 30 months would be in lower than the 5th percentile.
	ASQ				**NO** – Data are not reported for this milestone.

Note. CDC = Centers for Disease Control and Prevention; CDI = MacArthur–Bates Communication Development Inventories WG = Words and Gestures; WS = Words and Sentences.

[a]Age expected on the Ages and Stages Questionnaire–Third Edition (ASQ-3) was determined based on the age for the questionnaire on which the item was asked about; the latest age is presented if the item was asked about on multiple questionnaires; an additional screener (Parents' Evaluation of Developmental Status–Developmental Milestones) is included by Zubler et al. (2022) and not included in this illustrative example.

include perspectives from the necessary stakeholders (e.g., caregivers, speech language pathologists, and audiologists) to ensure that milestones are clinically applicable and accessible. The 15 months (i.e., tries to say one or two words besides mama or dada) and 18 months (i.e., tries to say ≥ 3 words besides mama or dada) milestones do not meet these criteria. Although the wording of the first two milestones reflect items on the ASQ-3 (the developmental screening tool that was referenced in the development of the CDC milestones), most of the literature quantifies children's expressive vocabulary as total number of words. Furthermore, the rationale to support the distinction between *mama* and *dada* from other words is unclear. For example, if one child has a total of four words, which are *mama*, *dada*, *ball*, *milk*, they do not meet the criteria for the 18 months milestone. Whereas if another child has a total of four words, which are *mama, ball, milk, dog,* they meet the criteria for the 18 months milestone. There are no data to suggest that the acquisition of the names of familiar caregivers is distinguishable from the acquisition of other early-acquired vocabulary words. Additionally, on the ASQ-3, caregiver names other than *mama* and *dada* are also included as "familiar caregivers." However, within the CDC milestone checklists, only *mama* and *dada* are listed, which may impact the inclusivity and accessibility of the milestones. Given that developmental surveillance systems are, in part, dependent on caregivers' identification of milestones, milestones must be easily identifiable, salient for families, and reflect the language environments of children in the United States. We recommend a more inclusive, clinically applicable milestone that represents the total number of words expected at a given age (e.g., tries to say 1–3 words).

Additionally, there is existing evidence (not included in the review by Zubler et al., 2022) that contradicts the new CDC expressive vocabulary milestones. In addition to the CDI data presented in the current article, Rescorla and Achenbach (2002) suggests that 75% of children produce 50 words between 21 and 23 months. Thus, there is no clear justification for moving the 50-words milestone from 24 months (as was the case in the previous version of the CDC milestones) to 30 months in the current version. Given that the number of words a child says is information that is easily collected and often reported to pediatricians, it is critical that the CDC update the new expressive vocabulary milestones to reflect the existing data. The existing data suggest that this milestone is met at the earliest 18 months and the latest 24 months. Given that CDC milestones are reflected at 18 month or 24 months (with no milestones reported between these two ages) and were designed to reduce the "wait and see" approach, the updated milestones should indicate that 75% of children should produce 50 words by 24 months of age, not 30 months of age.

These results also raise questions about the validity of the other CDC milestones and highlight the importance of transparency when making recommendations. For example, in addition to number of different words, the acquisition of two-word combinations by 24 months is an extensively researched milestones of expressive language and a strongly supported indicator of expressive language (Klee et al., 2000; Rescorla, 1989; Rescorla & Achenbach, 2002). Although this 24-month milestone remained unchanged, the revisions included an additional milestone at 30 months pertaining to the acquisition of two-word combinations that include action words. The development of this milestone suggests that it is necessary to distinguish the acquisition of two-word semantic combinations that include action words from the acquisition of two-word semantic combinations that include other word classes (e.g., nouns, adjectives, prepositions). The sources cited do not provide any evidence to support the decision to (a) distinguish the acquisition of two-word combinations with and without action words and (b) establish this milestone at 30 months (Accardo & Capute, 2005; Ertem et al., 2018; Lancaster et al., 2018; Tamis-Lemonda et al., 1998). In the absence of robust data, clinical experience may be necessary to establish milestones. However, professionals with clinical expertise in such areas (i.e., communication and hearing experts) were underrepresented in the development of the updated milestones. As such, it is critical for speech-language pathologists to establish a professional and advocacy role in the development and implementation of developmental surveillance initiatives.

More generally, the striking paucity of evidence supporting the new CDC-expressive vocabulary milestones illustrates the critical need for future research in this area to establish more accurate milestones for U.S. children, with a focus on collecting large-scale data that are inclusive of the United States' socioeconomic, racial/ethnic, linguistic, and cultural diversity. Although efforts are currently being made to increase the diversity of the CDI norming sample (Marchman et al., 2023), limitations with respect to racial and socioeconomic diversity remain, as the current CDI norming data includes overrepresentation of children who are White and from high socioeconomic backgrounds. One notable limitation is the absence of any normative data on language development for U.S. children who are bilingual or multilingual, even though more than a quarter of U.S. children hear a language other than English at home (Park et al., 2018). Vocabulary norms for bilingual children have been developed in other countries (see Floccia et al., 2018, for the United Kingdom and Singh et al., 2022, for Singapore), yet no such data are currently available for the United States. Indeed, although the CDC language/communication milestones were translated to multiple languages (CDC, 2022), the evidence

supporting these milestones did not include data from bilingual children and may not accurately reflect their language development (Hoff, 2023). It is beyond the scope of this article to discuss the extent to which the selected language/communication milestones do or do not apply to bilinguals, yet we urge caution when applying these milestones to populations not represented in the data. We also note the critical importance of updating the current milestones to include evidence from bilinguals, given the growing number of bilingual children in the United States and the lack of guidance for parents and pediatricians on how to identify signs of language and communication delays in these children (Glusman et al., 2021). In addition to linguistic diversity, other child, family, and environmental factors contribute to the variability in children's developmental trajectories. Therefore, across milestones and at different ages, it remains challenging to understand when children should be referred for further evaluation. Future work on establishing milestones may consider alternative methods to provide a more nuanced understanding of when to refer children for further evaluation. One particular approach is the use of item response theory, which establishes the probability of a child reaching a milestone given their developmental age as well as influential child, family, and environmental factors (Sheldrick et al., 2019). In doing so, providers may be better equipped to provide individualized recommendations, given each child's unique profile. Inspiration could also be drawn from the Global Scales for Early Development, a project led by the World Health Organization that aims to provide population level measures of child development with psychometrically stable performance across geographic, cultural, and language contexts (McCray et al., 2023). To develop these measures, quantitative and qualitative procedures were used to select theoretically relevant and globally feasible items representing child development for children across diverse cultural, demographic, social, and language contexts. It would be fruitful for similar efforts to be carried out in the United States.

Failure to identify children with potential language delays could have immediate consequences, given that early identification and provision of early intervention services is critical for children's long-term educational and mental health outcomes (Arkkila et al., 2008; Manning et al., 2019). The use of inaccurate data may influence resource allocation, such that crucial services are not provided to children who need them. Finally, this is an illustrative example that only addresses one type of language milestone, but it is possible that this misalignment with available data extends to milestones in other areas of development, such that relying on these milestones could result in widespread misidentification of children with developmental delays and disorders.

Data Availability Statement

All data extracted and interpreted for this review are included in this published article. Data collection tools are available from the first author upon request.

Acknowledgments

The authors received no external financial support for the research, authorship, and/or publication of this article.

References

Accardo, P. J., & Capute, A. J. (2005). *The Capute Scales: Cognitive Adaptive Test/Clinical Linguistic & Auditory Milestone Scale (CAT/CLAMS)*. Brookes. https://brookespublishing.com/wp-content/uploads/2020/10/Capute-Scales_Excerpt.pdf

American Community Survey. (2021). *Table B01001: Sex by age*. Census Reporter. https://censusreporter.org/tables/B01001/

American Speech-Language-Hearing Association. (2022). *ASHA statement on CDC's updated developmental milestones*. https://www.asha.org/about/statements/ASHA-Statement-on-CDC-Developmental-Milestones/

Arkkila, E., Rasanen, P., Roine, R. P., & Vilkman, E. (2008). Specific language impairment in childhood is associated with impaired mental and social well-being in adulthood. *Logopedics, Phoniatrics, Vocology, 33*(4), 179–189. https://doi.org/10.1080/14015430802088289

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest, 10*(3), 89–129. https://doi.org/10.1177/1529100610387084

Centers for Disease Control and Prevention. (2022). *CDC's developmental milestones*. https://www.cdc.gov/ncbddd/actearly/milestones/index.html

Dale, P. S., & Patterson, J. L. (2017). Early identification of language delay. In R. E. Tremblay, M. Boivin, R. D. e. V. Peters (Eds.), S. Rvachew (Topic Ed.), *Encyclopedia on early childhood development*. https://www.child-encyclopedia.com/language-development-and-literacy/according-experts/early-identification-language-delay

De Houwer, A. (2021). *Bilingual development in childhood*. Cambridge University Press. https://doi.org/10.1017/9781108866002

Ertem, I. O., Krishnamurthy, V., Mulaudzi, M. C., Sguassero, Y., Balta, H., Gulumser, O., Bilik, B., Srinivasan, R., Johnson, B., Gan, G., Calvocoressi, L., Shabanova, V., & Forsyth, B. W. C. (2018). Similarities and differences in child development from birth to age 3 years by sex and across four countries: A cross-sectional, observational study. *The Lancet Global Health, 6*(3), e279–e291. https://doi.org/10.1016/S2214-109X(18)30003-2

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., & Reilly, J. S. (1993). *The MacArthur Communicative Development Inventory: Words and sentences*. Singular.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur–Bates Communicative Development Inventories* (2nd ed.). Brookes.

Fischel, J. E., Whitehurst, G. J., Caulfield, M. B., & DeBaryshe, B. (1989). Language growth in children with expressive

language delay. *Pediatrics, 83*(2), 218–227. https://doi.org/10.1542/peds.83.2.218

Fisher, E. L. (2017). A systematic review and meta-analysis of predictors of expressive-language outcomes among late talkers. *Journal of Speech, Language, and Hearing Research, 60*(10), 2935–2948. https://doi.org/10.1044/2017_JSLHR-L-16-0310

Floccia, C., Sambrook, T. D., Luche, C. D., Kwok, R., Goslin, J., White, L., Cattani, A., Sullivan, E., Abbot-Smith, K., Krott, A., Mills, D., Rowland, C., Gervain, J., & Plunkett, K. (2018). I: Introduction. *Monographs of the Society for Research in Child Development, 83*(1), 7–29. https://doi.org/10.1111/mono.12348

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The Wordbank project*. MIT Press. https://doi.org/10.7551/mitpress/11577.001.0001

Genesee, F. (2008). Early dual language learning. *Zero to Three, 29*(1), 17–23.

Gilkerson, J., Richards, J. A., Warren, S. F., Oller, D. K., Russo, R., & Vohr, B. (2018). Language experience in the second year of life and language outcomes in late childhood. *Pediatrics, 142*(4), Article e20174276. https://doi.org/10.1542/peds.2017-4276

Gladstone, M., Lancaster, G. A., Umar, E., Nyirenda, M., Kayira, E., van den Broek, N. R., & Smyth, R. L. (2010). The Malawi Developmental Assessment Tool (MDAT): The creation, validation, and reliability of a tool to assess child development in rural African settings. *PLOS Medicine, 7*(5), Article e1000273. https://doi.org/10.1371/journal.pmed.1000273

Glusman, M., Pabalan, L., & Weisleder, A. (2021). Pediatrician and parent beliefs and practices about bilingual language development. *Pediatrics 147*(3_MeetingAbstract), 76. https://doi.org/10.1542/peds.147.3MA1.76a

Hammer, C. S., Morgan, P. L., Farkas, G., Hillemeier, M. M., Bitetti, D., & Maczuga, S. (2017). Late talkers: A population-based study of risk factors and school readiness consequences. *Journal of Speech, Language, and Hearing Research, 60*(3), 607–626. https://doi.org/10.1044/2016_JSLHR-L-15-0417

Hoff, E. (2023). Commonalities, differences, and differences that matter between monolingual and bilingual development. *Journal of Child Language, 50*(4), 818–822. https://doi.org/10.1017/S0305000922000770

Klee, T., Pearce, K., & Carson, D. K. (2000). Improving the positive predictive value of screening for developmental language disorder. *Journal of Speech, Language, and Hearing Research, 43*(4), 821–833. https://doi.org/10.1044/jslhr.4304.821

Lancaster, G. A., McCray, G., Kariger, P., Dua, T., Titman, A., Chandna, J., McCoy, D., Abubakar, A., Hamadani, J. D., Fink, G., Tofail, F., Gladstone, M., & Janus, M. (2018). Creation of the WHO Indicators of Infant and Young Child Development (IYCD): Metadata synthesis across 10 countries. *BMJ Global Health, 3*(5), Article e000747. https://doi.org/10.1136/bmjgh-2018-000747

Lipkin, P. H., Macias, M. M., Norwood, K. W., Jr., Brei, T. J., Davidson, L. F., Davis, B. E., Ellerbeck, K. A., Houtrow, A. J., Hyman, S. L., Kuo, D. Z., Noritz, G. H., Yin, L., Murphy, N. A., Levy, S. E., Weitzman, C. C., Bauer, N. S., Childers, D. O., Jr., Levine, J. M., Peralta-Carcelen, A. M., . . . Voigt, R. G. (2020). Promoting optimal development: Identifying infants and young children with developmental disorders through developmental surveillance and screening. *Pediatrics, 145*(1), Article e20193449. https://doi.org/10.1542/peds.2019-3449

Manning, B. L., Roberts, M. Y., Estabrook, R., Petitclerc, A., Burns, J. L., Briggs-Gowan, M., Wakshclag, L. S., & Norton, E. S. (2019). Relations between toddler expressive language and temper tantrums in a community sample. *Journal of Applied Developmental Psychology, 65,* Article 101070. https://doi.org/10.1016/j.appdev.2019.101070

Marchman, V. A., Dale, P. S., & Fenson, L. (2023). *The MacArthur–Bates Communicative Development Inventories: User's guide and technical manual* (3rd ed.). Brookes.

McCray, G., McCoy, D., Kariger, P., Janus, M., Black, M. M., Chang, S. M., Tofail, F., Eekhout, I., Waldman, M., van Buuren, S., Khanam, R., Sazawal, S., Nizar, A., Schönbeck, Y., Zongo, A., Brentani, A., Zhang, Y., Dua, T., Cavallera, V., . . . Gladstone, M. (2023). The creation of the Global Scales for Early Development (GSED) for children aged 0–3 years: Combining subject matter expert judgements with big data. *BMJ Global Health, 8*(1), Article e009827. https://doi.org/10.1136/bmjgh-2022-009827

Park, M., Zong, J., & Batalova, J. (2018). *Growing superdiversity among young US dual language learners and its implications*. Migration Policy Institute.

Paul, R. (1996). Clinical implications of the natural history of slow expressive language development. *American Journal of Speech-Language Pathology, 5*(2), 5–21. https://doi.org/10.1044/1058-0360.0502.05

Rescorla, L. (1989). The language development survey: A screening tool for delayed language in toddlers. *Journal of Speech and Hearing Disorders, 54*(4), 587–599. https://doi.org/10.1044/jshd.5404.587

Rescorla, L., & Achenbach, T. M. (2002). Use of the Language Development Survey (LDS) in a national probability sample of children 18 to 35 months old. *Journal of Speech, Language, and Hearing Research, 45*(4), 733–743. https://doi.org/10.1044/1092-4388(2002/059)

Sheldrick, R. C., & Perrin, E. C. (2013). Evidence-based milestones for surveillance of cognitive, language, and motor development. *Academic Pediatrics, 13*(6), 577–586. https://doi.org/10.1016/j.acap.2013.07.001

Sheldrick, R. C., Schlichting, L. E., Berger, B., Clyne, A., Ni, P., Perrin, E. C., & Vivier, P. M. (2019). Establishing new norms for developmental milestones. *Pediatrics, 144*(6), Article e20190374. https://doi.org/10.1542/peds.2019-0374

Singh, L., Cheng, Q., Tan, S. H., Tan, A., & Low, Y. L. (2022). Language acquisition in a multilingual society: English vocabulary norms and predictors in Singaporean children. *Child Development, 93*(1), 288–305. https://doi.org/10.1111/cdev.13676

Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaires, Third Edition (ASQ-3): A parent-completed child monitoring system*. Brookes.

Tamis-LeMonda, C. S., Bornstein, M. H., Kahana-Kalman, R., Baumwell, L., & Cyphers, L. (1998). Predicting variation in the timing of language milestones in the second year: An events history approach. *Journal of Child Language, 25*(3), 675–700. https://doi.org/10.1017/s0305000998003572

The Informed SLP Team. (2022). *No SLPs were in the room where it happened*. The Informed SLP, LLC. https://www.theinformedslp.com/review/no-sl-ps-were-in-the-room-where-it-happened

Visintainer, P. F., Leppert, M., Bennett, A., & Accardo, P. J. (2004). Standardization of the Capute Scales: Methods and results. *Journal of Child Neurology, 19*(12), 967–972. https://doi.org/10.1177/08830738040190121101

Zubler, J. M., Wiggins, L. D., Macias, M. M., Whitaker, T. M., Shaw, J. S., Squires, J. K., Pajek, J. A., Wolf, R. B., Slaughter, K. S., Broughton, A. S., Gerndt, K. L., Mlodoch, B. J., & Lipkin, P. H. (2022). Evidence-informed milestones for developmental surveillance tools. *Pediatrics, 149*(3), Article e2021052138. https://doi.org/10.1542/peds.2021-052138