Supplemental Material S1.

Additional participant enrollment

Several dyads were enrolled in the study pre-implantation, but did not continue with the study protocol and/or did not meet inclusion criteria, as follows. For *n* = 4 dyads, the mother was not a native speaker of English. For *n* = 11 dyads, the families were enrolled in the study pre-implantation and participated up through implantation, but elected not to complete any parts of the post-implantation protocol. For *n* = 12 dyads, the mother completed one or more recordings of her speech 3–15 months post-implantation, but the child did not contribute sufficient clinical speech-language data to construct a line of best fit on the Preschool Language Scales (PLS) total language score (Zimmerman et al., 2002), Peabody Picture Vocabulary Test–Fourth Edition (PPVT-4; Dunn & Dunn, 2007), or Reynell Developmental Language Scales (RDLS; Edwards et al., 1997) from post-implantation scores. Finally, for *n* = 2 dyads, there were missing recordings for the mother, i.e., no adult-directed (AD) condition recording and/or no spontaneous infant-directed (ID) condition recording at a post-implantation visit, due to equipment failure, experimenter error, and/or infant fussiness. Of dyads meeting inclusion criteria, Dyads 1–22 and 24 were enrolled in the first study protocol (2002–2008), while remaining dyads were enrolled in the second study protocol (2008–2013); see main text.

F0 measurement details

The F0 analysis procedure was designed to ensure accurate estimates of F0 from voiced portions of speech for each file. To achieve this goal, analysts first marked for exclusion any short stretches of voiced speech during the analyzed portions of recordings when mothers used non-modal voicing styles that involved disruption to vocal fold vibratory periodicity (i.e., diplophonia or creaky voice), which yield poor estimates of F0 (e.g., Talkin, 1995). Next, for the remaining portions of voiced speech, analysts inspected plots of F0 in Praat for the modal voiced portions using default F0 extraction parameters (minimum F0 = 75 Hz, maximum F0 = 600 Hz, silence threshold = 0.04, voicing threshold = 0.2) to assess the accuracy of F0 estimation through visual inspection of F0 curves combined with auditory impressions of pitch. If analysts determined that displayed default F0 estimates were incorrect, e.g., due to mother's use of a higher pitch than the default F0 maximum, then analysts changed values for up to four autocorrelation extraction parameters, cf. pitch floor (i.e., F0 minimum), pitch ceiling (i.e., F0 maximum), voicing threshold, and/or silence threshold, noting the required non-default parameter setting(s) necessary for that stretch of speech to show a visually accurate F0 display. If no altered parameter settings could be identified for a portion of speech that would yield an F0 curve on the display that was accurate, then that portion of speech was excluded. Finally, raw F0 values for each F0 epoch were then extracted using Praat's F0 autocorrelation algorithm using default F0 extraction parameter settings or else the custom parameter settings earlier identified by the analyst. Once the F0 values were extracted, any remaining outliers less than 75 Hz or greater than 900 Hz were removed before the median F0 pitch value for each mother was calculated in the ID and AD conditions. Note that nonparametric measures (median and inter-quartile range) were preferred for summary descriptive statistics of F0 values, based on prior findings showing distributions of most mothers' F0 values were positive skewed (Lennes et al., 2016).

Vowel Measurement Details

The vowel measurement procedure began by identifying all tokens of words with point vowels /i/, /ɑ/, and /u/ produced in extant spontaneous AD and ID recordings in first and second post-implantation intervals (i.e., the scheduled three and six months post-implantation intervals, respectively; see Table S2), including all spontaneous ID play condition recordings for dyads enrolled 2008–2013. To obtain adequate numbers of tokens, vowels from both content and function words were examined. Next, we identified tokens which met the following criteria: (a) the vowel token was at least 40 ms in duration; (b) the vowel was not followed by coda /r/ or /l/, which substantially changes pronunciations of preceding vowels; and (c) the vowel token must have been produced with full vowel quality, e.g., *you* spoken as /ju/ rather than schwa.

From the set of vowels meeting these criteria, phonetic analysts trained in formant analysis identified within each file the onset and offset of each randomly selected vowel token (see main text) via visual inspection of spectrograms and waveforms using segmentation criteria from the Buckeye Corpus (Pitt et al., 2007) within PRAAT 5.0.21 editor (Boersma & Weenink, 2017) software. Measurements of the first (F1) and second (F2) formants were then taken at the vowel midpoint using a combination of spectral slices, visual inspection of spectrograms, and linear predictive coding estimates; all F1 and F2 measurements were checked by hand to ensure correctness. Any tokens with a mean F0 of 350 Hz or higher were excluded for excessive measurement uncertainty associated with insufficient sampling of formant peaks due to wide harmonic spacing (Vallabha & Tuller, 2002). Further, any tokens for which the formants could not be reliably identified due to e.g., the mother talking too softly, overlapped child vocalization, background noise, etc. were examined by the first author, who made the final decision on whether the token was unusable. If a previously randomly selected token was discarded for any reason, it was replaced by another randomly selected token of that vowel from among the remaining tokens produced by that mother in the same speech condition.

To ensure reliable and valid measurement of the vowel space across ID and AD conditions, a criterion was set that F1 and F2 must have been successfully measured as above for a minimum of three tokens of each of /i/, /ɑ/, and /u/ in both ID and AD conditions. For five mothers (from Dyads 5, 7, 24, 25, and 35) the minimum three tokens of one or more corner vowels in ID and/or AD conditions could not be identified from recordings at the first two post-implantation intervals. For these mothers, further tokens of any corner vowels with fewer than three tokens in ID or AD conditions were therefore sought from additional spontaneous speech files at the third post-implantation interval (i.e., the scheduled 12-month post-implantation interval), and/or a pre-implantation interval (for the AD condition only). Upon consideration of an additional file for additional tokens of the corner vowel, tokens of that vowel were sought until all tokens of that vowel in the file had been measured and/or a maximum 18 tokens of that vowel were identified, whichever came first. After searching additional files, only one mother (from dyad 7) was excluded for insufficient vowel token counts.

Formant values in Hertz (Hz) were then converted to the mel scale as has been used in prior studies of vowel space (e.g., Bradlow, Torretta, & Pisoni, 1996; Englund & Behne, 2005; Kuhl et al., 1997; Lam & Kitamura, 2010, 2012) using the following formula (Fant, 1973, in Bradlow et al., 1996):

$$mels = \frac{1000}{\log 2} \times \log\left(\frac{Hz}{1000} + 1\right)$$

The mels conversion provided the basis for all analyses reported. The means and standard deviations of F1 and F2 were determined for each speaker in both ID and AD speech conditions. The formants of any vowel token whose F1 or F2 was two or more standard deviations away from a given participant's mean F1 or F2, respectively, were checked by hand to ensure accuracy. In addition, trained analysts re-measured a random selection of 5% of the tokens used in each speech sample for an analysis of inter-rater reliability. The percentage difference (Δ_i) between the first rater's measurement (r_1) and the second rater's measurement (r_2) was calculated using the equation (Kuhl et al., 1997):

$$\Delta_i = ((|\, r_1 - r_2 \,|) / r_1) \times 100\%$$

The average inter-rater percentage difference was 8.0% (*SD* = 8.4%). This is in line with reliability reported in previous studies and indicates high inter-rater reliability (e.g., Kuhl et al., 1997).

Vowel Space Area. Finally, vowel space triangles were constructed in an *x-y* plane, where the average *F1* and *F2* values of /i/, /ɑ/, and /u/ vowels were the respective *x* and *y* coordinates of the corners. The area of the resultant triangles in both ID and AD conditions was calculated using the following equation (Liu et al., 2003):

$$Area = \frac{|(F1_i * (F2_a - F2_u) + F1_a * (F2_u - F2_i) + F1_u * (F2_i - F2_a)|}{2}$$

Vowel Space Dispersion. Previous research on AD speech has identified vowel space dispersion as a good index of speech clarity (Bradlow et al., 1996). Vowel space dispersion is calculated by measuring the distance of each token from a central point in the talker's vowel space. This measure provides an indication of the overall expansion/compaction of the set of vowel tokens from each participant, and detects fine-grained individual differences in acoustic-phonetic characteristics (Bradlow et al., 1996). By capturing a slightly different aspect of vowel production characteristics than the traditional Heron method (Kuhl et al., 1997; Neel, 2008), this metric helps to provide an assessment of vowel clarity. Vowel space dispersion was calculated using the centroid of each speaker's vowel space triangle and averaging the distances of the individual tokens from the centroid (Bradlow et al., 1996) using the following equations. First, the centroid (C) of each speaker-condition vowel space triangle was calculated using the formula:

$$C = \left(\frac{F1_a + F1_i + F1_u}{3}, \frac{F2_a + F2_i + F2_u}{3}\right)$$

where /i/, /ɑ/, and /u/ were the corners of each vowel space triangle and *F1* and *F2* were the *x* and *y* coordinates of each of the corners. Next, the Euclidean distance ($|d|$) of each token from the centroid was calculated using the formula:

$$|d| = \sqrt{(F1_C - F1_t)^2 + (F2_C - F2_t)^2}$$

where $F1_C$ and $F2_C$ were the x and y coordinates of the centroid and $F1_t$ and $F2_t$ were the first and second formant values for the token in question. Finally, the vowel space dispersion (D) was calculated as the ratio of the Euclidean distances ($|d|$) of each token from the centroid of the triangle to the number of tokens (n), using the formula:

$$D = \frac{\sum|d|}{n}$$

Additional Analyses

We first considered the extent to which scores across child assessments were correlated with one another. Bivariate correlations among the four assessments were calculated for pairs of children contributing the minimum of two scores for each assessment. We report separately in Table S3 the bivariate correlations for two-year outcome scores (cf. intercepts) and for two-year change scores (cf. slopes). For outcome scores, two pairs were significantly correlated (out of six) – PLS vs. PPVT [$r(22) = .85$, $p < .01$] and RLDS-Expressive vs. RDLS-Receptive [$r(9) = .76$, $p < .05$] – while for change scores, one pair was significantly correlated (out of six) - RLDS-Expressive vs. RDLS-Receptive [$r(9) = .69$, $p < .05$].

Further, since maternal speech recordings were collected at variable time intervals from 3 to 12–15 months post-implantation, one possible concern is that variation ID speech might not reflect idiosyncratic individual differences among mothers, but rather systematic changes in speech stemming from longitudinal child development-related factors. To address these possibilities, we first calculated, for each mother, a (weighted) mean (in months post-implantation) of recording interval timing used in our analyses. To do so, we coded, at each possible recording interval (3-, 6-, or 12-months post-implantation), whether both ID and AD recording conditions were available at that interval for that mother (= 2), only one recording condition (= 1), or neither (= 0); the number of recordings for that mother was then summed to give a value $C_{\max}$. For example, the mother for Dyad 11 was missing an AD recording at 3 months post-implantation (= 1) but had both ID and AD recordings at 6 and 12 months post-implantation (= 2 for both); $C_{\max}$ was then 5 (= 1 + 2 + 2). The weighted mean of the timing of recordings for the mother was then (3 months * 1/5) + (6 months * 2/5) + (12 months * 2/5) = 7.8 months. A second method was also employed whereby we computed weighted means of timing of maternal *ID* recordings *only* following a similar approach but using a coding of 1 (= ID available) or 0 (= ID not available) for each post-implantation interval.

(i) *Mothers' recording timing vs. mothers' ID speech factors*. We first determined whether mothers' recording timing predicted values of their ID and AD speech variables in the models of child language attainment. Results revealed very little evidence of systematic relationships between maternal recording timing and maternal ID/AD speech variables. Variation in mothers' *lexical quantity* was not predicted by mean recording timing (ID and AD: $R = .08$, $p = .63$; ID only: $R = .06$, $p = .74$). Likewise, variation in mothers' *vowel space area* was not predicted by mean recording timing (ID and AD: $R = .26$, $p = .14$; ID only: $R = .30$, $p = .09$), nor was mothers' *vowel dispersion* (ID and AD: R

= .28, *p* = .10; ID only: *R* = .16, *p* = .35). There was a modest correlation between mothers’ *normalized F0 variability* and their mean recording timing (ID and AD: *R* = .37, *p* = .02; ID only: *R* = .42, *p* = .01). There was also a modest correlation between mothers’ *ID speech rate* and mean recording timing, but only when mothers’ mean recording timing based on both ID and AD recordings (ID and AD: *R* = .36, *p* = .04; ID only: *R* = .22, *p* = .21); since the significant correlation with ID rate was therefore driven by AD recording timing, this correlation must therefore be dismissed as spurious. These analyses suggest a lack of fundamental support for the proposition that variation in mothers’ mean post-implantation recording timing can account for variation in their ID speech properties—where the lack of correlation between recording timing and the two speech factors that emerged most often significant from child language attainment modeling (i.e., lexical quantity and vowel space area differences) is particularly striking.

(ii) *Mothers’ recording timing vs. child age at implantation.* Next, we examined whether any of several longitudinal child development-related factors were predicted by was systematically related to the child’s age of implantation. Children’s age of implantation was not significantly predicted by the weighted mean timing estimates of maternal recordings. This was true whether timing estimates were based on both ID and AD recordings (*R* = .13, *p* = .45) or only ID recordings (*R* = .11, *p* = .53).

(iii) *Mothers’ recording timing vs. timing of child language assessments.* Next, we considered whether maternal recording timing was systematically related to the timing of child language assessment administration. To address this, we calculated for each child included in language outcome models for a given assessment, the mean post-implantation timing of that assessment (see Table S4). We found almost no relationship between maternal recording timing and evidence of systematic relationships between maternal recording timing and child assessment timing, either for PLS (both ID and AD: *R* = .06, *p* = .72; ID only: *R* = .20, *p* = .27), PPVT (both ID and AD: *R* = 029, *p* = .16; ID only: *R* = .46, *p* = .03), or RDLS-Expressive and RDLS-Receptive (both ID and AD: *R* = 024, *p* = .50; ID only: *R* = .21, *p* = .56).

(iv) *Child language growth estimates vs. post-implantation time elapsed under an exponential language growth function*. Finally, we considered whether our interpretation of our conclusions based on linear modeling of assessment data might be undermined by the possibility that language growth curves may have been exponential in character. Recall that the linear regression modeling approach was necessitated by sparse data for some children; however, given this approach, regression line slopes were the basis of calculating estimates of meaningful individual differences in language growth for each child over a two-year interval (i.e., the “rise” calculated over the modeled two-year “run”). However, putative exponential language growth patterns in time imply systematic, monotonic changes to the size of the “rise” calculated from a linear estimate of growth over a fixed time interval as developmental time elapses. Since in the present study children were administered assessments at differing times (on average) and over variable timespans, the possibility therefore obtains that calculated slope differences interpreted by us to reflect meaningful individual variation in language attainment actually trivially reflect similar exponential patterns of growth that readily account for these slope differences.

To test this possibility against our data, we conducted an analysis to determine whether slopes of regression lines taken as indices of individual differences in child language growth were systematically related to the (mean) timing of post-implantation assessments. If exponential growth curves can explain measured slope differences, we should see shallower slopes when assessments were administered earlier, vs. steeper slopes when they were administered later (or vice versa). We focused on PLS and PPVT scores to test this exponential growth explanation, since the small number of children assessed on the RDLS were expected to result in an underpowered test of the hypothesis. Note, however, that relating linear regression slopes to putative underlying exponential growth curves is complicated by attested variability in the timespans over which assessments were administered to the children. To control for this variability, we therefore identified those children who had been administered PLS and PPVT assessments over a relatively "narrow" range of time intervals, taken here to be 36 months or less from the first to the last administration. This yielded a subset of $N = 23$ and $N = 14$ children for the PLS and PPVT, respectively. For these children, we then computed the (weighted) mean timing (post-implantation) of assessment administration timing for the PLS or PPVT and subsequently divided each data subset into relatively "early" vs. "late" assessment administration groups by performing a median split within groups. Finally, we tested whether "early" and "late" administration of assessments across children systematically and reliably predicted differences in slope values, as predicted under an exponential growth explanation. Using a series of two-tailed, independent samples *t*-tests, we found no significant differences in language growth estimates for early vs. late assessment administration for the PLS [$t(21) = 1.48$, $p = .15$], nor for the PPVT [$t(12) = 1.95$, $p = .08$]. ("Early" and "late" administration timing also did not map to systematic differences in two-year outcome score estimates for the PLS [$t(21) = .51$, $p = .62$] nor for the PPVT [$t(12) = 1.43$, $p = .18$].) In summary, no support was found for a hypothesis that language growth estimates derived from linear regressions could be explained in terms of similar language growth trajectories across children that were putatively exponential in nature, coupled with confounded variability in assessment administration timing. While it is true that some, but not other, aspects of language growth *may* be more accurately characterized as exponential, rather than linear, functions (Huttenlocher et al., 2010), the present analyses support the validity of the methods used for modeling changes in language assessment values over time as mapping to true individual differences in language growth. Importantly, these differences in modeled language growth over children were crucially more reliably and strongly ($R > .5$) predicted statistically by individual differences in maternal ID speech factors, than any indices of developmental timing or other extraneous variables examined here.

Table S1. Child hearing-assistive device type and processor, and deafness etiology for children with cochlear implants in the present study.

ID	Child Age at Implantation (months)	Device	Processor	Deafness Etiology
1	12.7	Nucleus 24 K	Sprint	genetic
2	13.8	Nucleus 24 Contour	Sprint	auditory neuropathology
3	11.8	Nucleus 24 Contour	Tempo +	genetic
4	10.3	Med-El C 40+	Sprint/Freedom	Waardenburg syndrome
5	22.5	Nucleus 24 Contour	Sprint	unknown
6	24.2	Nucleus 24 Contour	Sprint	unknown
7	16.1	Nucleus 24 K	Sprint/Freedom	unknown
8	16.8	Nucleus 24 Contour	Tempo + (L) PSP ®	unknown
9	16.5	Med-El C 40+ (L)	Sprint/Freedom	unknown, genetic
10	24.2	Nucleus 24 Contour	Sprint	unknown
11	8.3	Nucleus 24 Contour	Sprint/Freedom	genetic
12	10.4	Nucleus 24 Contour	Freedom	auditory neuropathy
13	16.7	Nucleus Freedom – Contour Advance	PSP/Harmony	Mondini, premature
14	21.5	Clarion HiRes 90k	Freedom	unknown
15	17.9	Nucleus Freedom – Straight	Freedom	auditory neuropathy, bacterial meningitis
16	13.2	Nucleus Freedom – Contour Advance	Freedom	unknown
17	12.8	Nucleus Freedom – Contour Advance	Freedom	unknown
18	22.7	Nucleus Freedom – Contour Advance	Freedom	unknown
19	10.2	Nucleus Freedom – Contour Advance	Freedom	genetic
20	11.9	Nucleus Freedom – Contour Advance	PSP/Harmony	unknown, premature
21	24.2	Nucleus Freedom – Contour Advance	Freedom	unknown, premature
22	10.3	Nucleus Freedom – Contour Advance	Freedom	unknown
23	16.0	Nucleus Freedom – Contour Advance	Freedom	unknown, genetic
24	9.9	HiRes 90k Focus	Freedom	unknown
25	13.7	Nucleus Freedom – Contour Advance	Freedom	unknown
26	21.9	Nucleus Freedom – Contour Advance	Freedom	unknown
27	15.9	Nucleus Freedom – Contour Advance	Freedom	unknown
28	13.7	Nucleus Freedom – Contour Advance	Unknown	Mondini
29	9.8	Unknown	Unknown	Connexin 26
30	14.5	Nucleus Freedom System 5	Freedom	unknown
31	8.2	Nucleus 5	CP800	unknown
32	9.0	Nucleus System 5	CP810	Waardenburg syndrome
33	16.8	Nucleus C512	Contour advance	unknown
34	16.3	Nucleus C5	CP810	genetic
35	13.5	Nucleus C512	Unknown	auditory neuropathy
36	22.9	Unknown	Unknown	auditory neuropathy

Table S2. Recording sessions attended by mothers for protocol visits scheduled three, six, and 12 months after the child's CI surgery, along with weighted mean of recording interval timing for both ID and AD recordings, or only ID recordings. (See Supplemental Analyses.) X = extant recording session for the scheduled post-implantation interval. n/a = no recordings available for the interval. [a]No AD spontaneous recording available for the interval. [b]No ID spontaneous recording available for the interval. [c]Lab visit took place 2-3 months after scheduled interval.

ID	Recording interval (in months post-implantation)			Weighted mean of recording interval timing (in months post-implantation)	
	3	6	12	Both ID and AD	ID only
1	X	X	X	7	7
2	X	X	X	7	7
3	X	X	X	7	7
4	X	X	X	7	7
5	n/a	n/a	X	12	12
6	X	X	n/a	4.5	4.5
7	X	X	X	7	7
8	X	X	X	7	7
9	X	X	X	7	7
10	n/a	X[c]	n/a	8	8
11	X[a]	X	X	7.8	7
12	X	X	n/a	4.5	4.5
13	X	X	X	7	7
14	X	X	X	7	7
15	X	n/a	X	7.5	7.5
16	X	X	X	7	7
17	X	X	X	7	7
18	X	X	X[a]	6	4.5
19	X	X	n/a	4.5	4.5
20	X	X	X	7	7
21	X	X	X	7	7
22	X	X	X	7	7
23	X	X	X	7	7
24	n/a	X[a]	X[a]	9	9
25	X	X[a]	X	7.2	7
26	X	n/a	n/a	3	3
27	X[a]	X[b]	X[b,c]	8	3
28	X	X	X	7	7
29	X	X	n/a	4.5	4.5
30	X	X	X	7	7
31	X	X	X	7	7
32	X	X	X	7	7
33	X	X	n/a	4.5	4.5
34	n/a	X	X[c]	15	10.5
35	X	X	n/a	4.5	4.5
36	n/a	X	X[c]	10.5	10.5

Table S3. (a) Bivariate correlations across children for pairs of assessments for outcome scores. (b) Bivariate correlations across children for pairs of assessments for change scores. Degrees of freedom are shown in parentheses. **p* < .05 ***p* < .01.

(a)	PLS	PPVT	RDLS-Expressive	RDLS-Receptive
PLS	1.0			
PPVT	.85** (22)	1.0		
RDLS-Expressive	.46 (6)	.36 (7)	1.0	
RDLS-Receptive	.36 (6)	.44 (7)	.76* (9)	1.0

(b)	PLS	PPVT	RDLS-Expressive	RDLS-Receptive
PLS	1.0			
PPVT	.26 (22)	1.0		
RDLS-Expressive	.44 (6)	.06 (7)	1.0	
RDLS-Receptive	.74 (6)	-.09 (7)	.69* (9)	1.0

Table S4. Longitudinal assessments administered to each child, shown as ordinal numbers of 6-month post-implantation increments (e.g., 1 = 6 months post-implantation, 2 = 12 months post-implantation, etc.) Values with asterisks correspond to singleton administrations not meeting minimum data requirements for linear regression. Values in square brackets give cell means; the bottom row gives the means of bracketed child cell means in each column. n/a = not applicable/no assessment.

	Post-implantation interval (number of 6-month increments)			
ID	**Any assessment**	**PLS**	**PPVT**	**RDLS**
1	3,4,5,7,8,10,11,12,17 [8.6]	8,10 [9.0]	10,11,12,17 [12.5]	3,4,5,7 [4.8]
2	3,4,5,6,7,9,10,12,14,16 [8.6]	n/a	9,10,12,14,16 [12.2]	3,4,5,6,7 [5.0]
3	3,4,5,6,7,8,9,10,12 [7.1]	8,9 [8.5]	10,12 [11.0]	3,4,5,6,7 [5.0]
4	2,3,4,5,6,8,9 [5.3]	8,9 [8.5]	9*	2,3,4,5,6 [4.0]
5	3,4,9,10,12,14,18 [10.0]	n/a	9,10,12,14,18 [12.6]	3,4 [3.5]
6	2,3,4,5,6 [4.0]	n/a	n/a	2,3,4,5,6 [4.0]
7	4,5,6,7,8,10,12 [7.4]	7,8,10 [8.3]	8,10,12 [10.0]	4,5,6 [5.0]
8	1,2,3,4,6,7,8,9 [5.0]	7,8 [7.5]	8,9 [8.5]	1,2,3,4,6 [3.2]
9	2,4,6,7 [4.8]	6,7 [6.5]	6,7 [6.5]	2,4 [3.0]
10	4,5,6,7,8,12,14 [8.0]	5,7 [6.0]	6,7,8,12,14 [9.4]	4*
11	3,4,5,6,7,8,11,14,18 [8.4]	5,6,8 [6.3]	6,7,8,11,14,18 [10.7]	3,4 [3.5]
12	2,3,4,5,6,8,10 [5.4]	3,4,5,6,8 [5.2]	5,6,8,10 [7.3]	2*
13	3,5,6,7,8 [5.8]	3,5,6,7,8 [5.8]	n/a	n/a
14	4,5,15 [8.0]	4,5 [4.5]	5,15 [10.0]	n/a
15	1,2,3,4,5,6,7,8,9,11,14,15 [7.1]	2,3,4,5,6,7 [4.5]	3,4,5,6,7,8,9,11,14,15 [8.2]	1*
16	2,3 [2.5]	2,3 [2.5]	n/a	n/a
17	2,3,4,19 [7.0]	2,3,4 [3.0]	4,19 [11.5]	n/a
18	1,2,4,6,7,10 [5.0]	1,2,4,6,7 [4.0]	2,4,6,7,10 [5.8]	n/a
19	1,3,5,6,7 [4.4]	1,3,5,7 [4.0]	6*	n/a
20	1,2,3,4,6 [3.2]	1,2,3,4,6 [3.2]	4,6 [5.0]	n/a
21	1,2,3,4,5,6,7,10 [4.8]	1,2,3,4,5,6,7,10 [4.8]	3,4,5,6,7,10 [5.8]	n/a
22	1,2,3,4 [2.5]	1,2,3,4 [2.5]	4*	n/a
23	1,2,3,4 [2.5]	1,2,3,4 [2.5]	4*	n/a
24	1,3 [2.0]	1,3 [2.0]	n/a	n/a
25	1,2,3,4,5,8,10,11,13 [6.3]	1,2,3,4,5,8,10 [4.7]	5,8,10,11,13 [9.4]	n/a
26	1,2 [1.5]	1,2 [1.5]	n/a	n/a
27	1,2 [1.5]	1,2 [1.5]	n/a	n/a
28	1,2,6,8,10 [5.4]	1,2,6,8,10 [5.4]	6,8,10 [8.0]	n/a
29	1,2,3,4,5,6,7,10 [4.8]	1,2,3,6,7,10 [4.8]	4,5,7,10 [6.5]	n/a
30	1,2,4,6,8,9 [5.0]	1,2,4,6,8,9 [5.0]	4,6,8,9 [6.8]	n/a
31	1,2,4,7,8 [4.4]	1,2,4,7,8 [4.4]	4,7,8 [6.3]	n/a
32	1,2,4,6,8 [4.2]	1,2,4,6,8 [4.2]	4,6,8 [6.0]	n/a
33	1,2,5,8 [4.0]	1,2,5,8 [4.0]	2,5,8 [5.0]	n/a
34	1,6,9 [5.3]	1,6,9 [5.3]	6,9 [7.5]	n/a
35	1,2,4,6,8,10 [5.2]	1,2,4,6,8 [4.2]	6,8,10 [8.0]	n/a
36	1,2,5 [2.7]	1,2,5 [2.7]	n/a	n/a
M	5.2 (2.6 years)	4.8 (2.4 yrs)	8.4 (4.2 yrs)	4.1 (2.0 yrs)

References

Boersma, D. C., & Weenink, D. (2017). *Praat: Doing phonetics by computer* (Version 6.0.29). https://www.praat.org

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic–phonetic talker characteristics. *Speech Communication, 20*(3), 255–272. https://doi.org/10.1016/S0167-6393(96)00063-5

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4: Peabody Picture Vocabulary Test–Fourth Edition*. Pearson Assessments. https://doi.org/10.1037/t15144-000

Edwards, S., Fletcher, P., Garman, M. A. H., Letts, C., & Sinka, I. (1997). *Reynell Developmental Language Scales III: The University of Reading Edition*. Western Psychological Services.

Englund, K. T., & Behne, D. M. (2005). Infant directed speech in natural interaction—Norwegian vowel quantity and quality. *Journal of Psycholinguistic Research, 34,* 259–280.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. Cognitive Psychology, 61(4), 343–365. https://doi.org/10.1016/j.cogpsych.2010.08.002

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*(5326), 684–686. https://doi.org/10.1126/science.277.5326.684

Lam, C., & Kitamura, C. (2010). Maternal interactions with a hearing and hearing-impaired twin: Similarities and differences in speech input, interaction quality, and word production. *Journal of Speech, Language, and Hearing Research, 53*(3), 543–555. https://doi.org/10.1044/1092-4388(2010/09-0126)

Lam, C., & Kitamura, C. (2012). Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation. *Developmental Science, 15*(2), 212–221. https://doi.org/10.1111/j.1467-7687.2011.01118.x

Lennes, M., Stevanovic, M., Aalto, D., & Palo, P. (2016). Comparing pitch distributions using Praat And R. *The Phonetician, 111-112,* 9–34.

Liu, H.-M., Kuhl, P. K., & Tsao, F.-M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science, 6*(3), F1–F10. https://doi.org/10.1111/1467-7687.00275

Neel, A. T. (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research, 51*(3), 574–585. https://doi.org/10.1044/1092-4388(2008/041)

Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. (2007). *Buckeye Corpus of Conversational Speech* (2nd release). Department of Psychology, Ohio State University. http://www.buckeyecorpus.osu.edu

Talkin, D. (1995). A robust algorithm for pitch tracking (RAPT). In W. B. Kleijn & K. K. Paliwal (Eds.), *Speech coding and synthesis* (pp. 495–518). Elsevier Science.

Vallabha, G. K. & Tuller, B. (2002). Systematic errors in the formant analysis of steady-state vowels. *Speech Communication, 38*(1–2), 141–160. https://doi.org/10.1016/S0167-6393(01)00049-8

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool Language Scale–Fourth Edition*. Pearson.