Published in final edited form as:

Int J Pediatr Otorhinolaryngol. 2009 April ; 73(4): 561–571. doi:10.1016/j.ijporl.2008.12.005.

A comparison of a child's fundamental frequencies in structured elicited vocalizations versus unstructured natural vocalizations: A case study

Eric J. Hunter

National Center for Voice and Speech, Denver, CO

Abstract

Objectives—Building on the concept that task type may influence fundamental frequency (F_0) values, the purpose of this case study was to investigate the difference in a child's F_0 during structured, elicited tasks and long-term, unstructured activities. It also explores the possibility that the distribution in children's F_0 may make the standard statistical measures of mean and standard deviation less than ideal metrics.

Methods—A healthy male child (5 years, 7 months) was evaluated. The child completed four voice tasks used in a previous study of the influence of task type on F_0 values: (1) sustaining the vowel /a/; (2) sustaining the vowel, /a/, embedded in a word at the end of a phrase; (3) repeating a sentence; and (4) counting from 1 to 10. The child also wore a National Center for Voice and Speech voice dosimeter, a device that collects voice data over the course of an entire day, during all activities for 34 hours over 4 days.

Results—Throughout the structured vocal tasks within the clinical environment, the child's F_0 , as measured by both the dosimeter and acoustic analysis of microphone data, was similar for all four tasks, with the counting task the most dissimilar. The mean F_0 (~257 Hz) matched very closely to the average task results in the literature given for the child's age group. However, the child's mean fundamental frequency during the unstructured activities was significantly higher (~376 Hz). Finally, the mode and median of the structured vocal tasks were respectively 260 Hz and 259 Hz (both near the mean), while the unstructured mode and median were respectively 290 Hz and 355 Hz.

Conclusions—The results of this study suggest that children may produce a notably different voice pattern during clinical observations compared to routine daily activities. In addition, the child's long-term F_0 distribution is not normal. If this distribution is consistent in long-term, unstructured natural vocalization patterns of children, statistical mean would not be a valid measure. Mode and median are suggested as two parameters which convey more accurate information about typical F_0 usage. Finally, future research avenues, including further exploration of how children may adapt their F_0 to various environments, conversation partners, and activity, are suggested.

Keywords

child; acoustic analysis; fundamental frequency; voice disorders

Reprint Request and Corresponding Author: Eric Hunter, 1101 13th Street, Denver, CO 80204, Tel: 303-446-4839, Fax: 303-893-6487, E-mail: ehunter@dcpa.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Acoustic analysis of voice and speech is a non-invasive, quantifiable method that is frequently used as part of a voice evaluation of normal and disordered voices. Several studies have attempted to translate these acoustic analyses to children's voices (e.g., [1–4]). However, usually these studies focus on a few measures rather than the wide range typically collected for adult voices. Further, many such studies have focused on older, rather than younger preschool/kindergarten-age children. One reason for adopting such an approach might be that obtaining valid acoustical data from young children is hampered by the difficulty in giving appropriate instruction and ensuring compliance. For example, in their voice range profiles of 277 children (range 5–14 years), Böhme and Stuchlik [5] suggested that this metric was not valid for children under 7 as these young children primarily produced only a single tone or two regardless of the prompted pitch target.

Nevertheless, despite these difficulties, an acoustic assessment is an important part of quantifying pediatric voice use, misuse, and disorders. One particular acoustic metric commonly used is vocal fundamental frequency (F_0), which can be extracted from a variety of transducers attached to the subject during a range of speech tasks (e.g., [6,7]). Table 1 contains a synopsis of F_0 studies, found through an extensive search of PubMed, examining native U.S. English-speaking children aged 2–7 years. Tasks may include imitating a specific phoneme (e.g., [8]); reading or repeating words, phrases, sentences or paragraphs (e.g., [9, 10]); naming or describing pictures (e.g., [6]); or responding to uniform questions (e.g., [11]).

One weakness apparent in the body of child F_0 literature is the use of incomplete, inconsistent and often unclear data reporting metrics. Most studies provide mean F_0 and standard deviation. These are usually reported as the mean of the subjects' means and standard deviation of the subjects' means, but rarely is the actual mean F_0 or standard deviation for each subject or the mean standard deviation reported. Thus, it is difficult to conclude whether these calculations represent inter-subject or intra-subject variability. Further, while studies report F_0 range as the range of mean F_0 values across subjects, few report the actual overall F_0 range across subjects (i.e., the maximum and minimum, or $F_{0\text{ max}}$ and $F_{0\text{ min}}$, the subjects can produce). This weakness is exacerbated by the fact that it is not always specified clearly in what way the mean, standard deviation, and range are being reported.

A second weakness is the variability in the type of tasks used to elicit vocalizations in children, which at least one recent study suggests may affect the F_0 actually produced. In their study published recently in this journal comparing the validity and variability of four different clinical protocols for F_0 measurements, Baker et al. [12] determined that the type of task used elicited significantly different results. Specifically, they found that children's F_0 produced during a counting task was significantly greater than that produced during a phrase or sentence task. They postulate that this higher F_0 might be affected by the ability of the children to produce this task with less instruction from the investigator (thus, perhaps, with less elicitation effect).

A final possible weakness is that child F_0 studies rarely examine long-term, natural unstructured vocalizations. If vocal task type used to examine F_0 in children affects the data collected within the clinic, it is also likely that F_0 collected in the clinic might differ from F_0 produced during natural vocalizations outside of the clinic. Such differences would be useful to know. Two studies appear to point to the possible difference between F_0 in structured versus unstructured speech. In their comparison of the F_0 in stuttering versus non-stuttering preschool-aged boys, Hall and Yairi [7] obtained speech samples from elicited, unstructured conversations between each boy and his parent, as well as a brief conversation between each boy and one of the investigators (minimum of 70 utterances per child). The mean F_0 of the control group of 10 non-stutterers was 310 Hz (mean 46 months old, or almost 4 y/o), similar

to the mean F_0 of children (25 months old, or just over 2 y/o) calculated from elicited free speech using familiar books and toys [13]. It is important to note that both studies obtain speech samples from relatively spontaneous, unstructured speech (although the conversations were elicited within a clinical environment, rather than in normal daily speech). It is important to note that the reported results from Hall and Yairi are quite dissimilar from children closer in age whose F_0 was calculated from elicited, structured speech (e.g., [8–10,14]).

Further, few studies collect extreme speech (loud and/or high pitch, or very excited speech) or non-speech vocalizations from children, both of which would provide a more complete picture of how children use their voices in their natural environments. One study was found which examined laughter in young children ($n = 4$; age = 3) [15]. In their study, laughter was broken into 7 different types of laughter with different communicative intent and F_0 values: [1] dull comment ($F_{0\text{mean}}=395.0$ Hz); [2] exclamatory comment ($F_{0\text{mean}} = 416.0$ Hz); [3] chuckle ($F_{0\text{mean}}=439.4$ Hz); [4] basic rhythmical ($F_{0\text{mean}}= 469.5$ Hz); [5] variable rhythmical ($F_{0\text{mean}}=410.8$ Hz); [6] classical rhythmical ($F_{0\text{mean}}=494.4$ Hz); and [7] squeal ($F_{0\text{mean}}=1906.9$ Hz).

Unfortunately, no study was found that collected unstructured F_0 data from vocalizations collected in a child's natural environment, as opposed to a clinical or other very controlled setting. Further, no study was found which evaluated long-term F_0 in children over many hours or days. Finally, with the exception of F_0 studies of infants/pre-talkers (e.g., [16–18]), few studies of young children attempt to evaluate F_0 from purely unstructured (i.e., non-elicited, spontaneous) speech, capturing a combination of speech and non-speech vocalizations.

These gaps in the current body of research reduce our ability to truly understand a child's voice use, including the full spectrum of vocalizations which might not only trigger vocal health disorders such as pediatric vocal nodules, but also provide the symptomatic markers on the voice of such disorders. One study [19] explored the incidence of vocal nodules among school-age children (7–16 y/o). Using acoustic analysis and rigid stroboscopy, they found that 30.3 percent of the 617 children had some degree of vocal nodules: 13.3%, minimal lesion; 14.3%, immature nodule; 2.6%, mature nodule; and 0.2%, vocal polyp. An additional study of 646 patients at a voice center found this condition in 40 percent of children [20] and primarily among male children. If studies incorporate acoustic analyses into diagnosis and monitoring of such conditions, it would seem crucial that these data gaps be filled.

To fill these data gaps, it would be useful to be able to unobtrusively examine a child's F_0 from all vocalizations during normal daily activities over a long period of time. A potential tool for such examination is the National Center for Voice and Speech (NCVS) voice dosimeter (Figure 1, discussed in further detail below), a device designed to monitor voice use in adults for several days at a time ([21–23]). One of the parameters captured by the device and used in dose calculation is F_0 . Because the transducer is an accelerometer, only skin vibration (primarily from the vocal folds) is transduced and airborne sound waves (and, thus, actually speech and language) are not. Previous studies have used the voice dosimeter to measure long-term voice use in classical singers and teachers (e.g., [24,25]). Subjects in these studies typically wore the devices 18 hours per day for two weeks.

As discussed previously, it is possible that our picture of a child's F_0 , as calculated in previous studies by traditional clinical tasks, is not an adequate or complete representation of what a child is actually producing during normal daily activities. The current study explores potential gaps through a case study by comparing an individual child's F_0 during structured, elicited vocalizations to *his own* F_0 during long-term, unstructured, natural vocalizations. Two research questions are asked. First, does a child produce similar F_0 patterns during structured vocal tasks in a clinical evaluative environment (i.e., elicited responses and few non-speech vocalizations),

as during long-term unstructured activities (i.e., both speech and non-speech vocalizations)? Second, does the F_0 during these activities follow a normal distribution, thereby allowing usual basic statistical metrics (i.e., mean and standard deviation) to be properly used?

2. Methods

As part of the research team working with the NCVS voice dosimeter, the author was testing a new version of the dosimeter software by wearing the device during nearly all waking hours over a 6-week interval. During this time, the author's young son and daughter repeatedly requested to wear the device. After attaching the device, the author's daughter asked to have it removed almost immediately. The son wanted to keep it on and, thus, later became the subject for the current case study (described below).

2.1 Participant

The young boy (age=5 years, 7 months) was a native English speaker. He passed basic hearing screenings conducted by an audiologist and was observed by two speech pathologists to have no resonance disorders or voice disturbances. The child was engaged in age-appropriate educational programs, namely thrice weekly preschool classes during the recording period. He was scheduled to begin year-round kindergarten two months after the recordings reported here were taken.

2.2 Procedure

The procedures for the current study consisted of two parts (not in time-sequenced order). Part one included a short set of vocal tasks produced in a controlled acoustic environment based on the study protocol used by Baker et al. [12], briefly described below. Part two used the NCVS voice dosimeter to capture the child's long-term, unstructured F_0 use from speech and non-speech vocalizations in a variety of natural daily activities.

2.2.1. Structured Vocal Tasks in a Controlled Acoustic Environment—Four specific vocal tasks taken from Baker et al. [12] were used to elicit vocal fold fundamental frequency differences. We chose to use this study, not only because of the diversity of elicited tasks used, but also to test whether the child was representative of a larger set of similar-aged children during structured vocal tasks. Briefly, the tasks were (with italicized labels in parentheses): [1] a sustained /a/ for 5 s (*ahhh*); [2] the sentence "I need a mop" in which the child was asked to sustain the /a/ for 5 sec (the investigator's hand displayed the duration for the child) and not say the /p/ at the end of the word (*mop*); [3] the sentence, "Bob wants a ball" (*ball*); and [4] a count from 1–10 at a comfortable rate of speed (*count*). Each of the four tasks was repeated 3 times, and the order was randomly presented via cue cards. The parent/author, who was a regular presence during the unstructured voice dosimeter portion of the study, was chosen to be the elicitor. This decision was made in order to reduce an additional variable when comparing structured and unstructured vocalizations (i.e., an unfamiliar elicitor causing a difference in F_0). To counteract potential bias, the elicitation protocol presented in Baker et al. was strictly followed during the recording session.

Recordings were conducted in a single-wall IAC sound isolation booth (~10ft × 10ft). The voice dosimeter was worn by the child for the entire in-booth recording protocol. In addition, the child wore a head-mounted microphone (Countryman Associates omnidirectional B3 Lavalier, less than ±1 dB over the entire voice range) custom-mounted on a wire boom attached to a plastic frame, worn like a pair of glasses. The microphone element was about 5 cm from the mouth and slightly to the side, out of the airstream. A Type I sound level meter (Brüel & Kjaer 2238 Mediator) set at linear frequency weighting was mounted on a stand with a 30 cm string attached for quick measurements during initial calibration of the head-mounted

microphone. Because of the movements of the child, repeated measurements were made to ensure that a consistent distance during the calibration procedure was maintained (see Svec et al. [21], for details). After the calibration procedure, the sound level meter was retracted to simply act as a chatter channel (ambient discussion from investigator and subject) in the recording. Although the child sat during the recording session, the setup was designed so that the child was able to stand up and move around after the initial calibration if he desired.

The head-mounted microphone and sound level meter were connected to a Millennia HV-3D 8 channel microphone pre-amp (Millennia Media). The pre-amp delivered the signal to an RME ADI-8 DS AD/DA Converter and an RME Multiface II 36 channel digital audio interface (RME, Germany). The digital interface to the recording computer was via a proprietary PCI expansion card on the motherboard. All signals were recorded on a PC at a 44.1 kHz sampling rate and 24-bit accuracy using CUBASE SE software (version 3.0.3) as the software interface.

2.2.2. Extended Dosimeter Study of Unstructured Vocalizations—The NCVS voice dosimeter measures vocal dose [26] by an accelerometer attached with surgical glue near the sternal notch (Figure 1, left). It captures raw acceleration data at a sampling frequency of 11,025 Hz and processes the data in 30 ms intervals in real time. Calculations of F_0 (Hz) and other acoustic measurements are then stored every 30 ms. The dosimeter can collect more than 70 hours of voice data (in these 30 ms steps) and about 25 hours on one battery (hot swappable).

The child wore the NCVS voice dosimeter (Figure 2) for 3 days during one week. The device itself was placed in a small backpack so that the child could have nearly full freedom of motion without hindrance. The three days were chosen to reflect a variety of normal activities for the child. Day one (10.5 total recording hours) consisted of 8.5 hours at home with family (2 parents and 2 younger female siblings) and 2 hours at a peer's birthday party (much of it playing in a gymnasium). Day two (11 hours total recording) included 8 hours at home with family and 3 hours at church (1 hour relatively quiet during a congregational meeting and 2 hours with similar-aged children during a children's meeting with singing and instruction). Day three (7 total usable hours) consisted of 3.5 hours at home with one parent and both siblings and 3.5 hours of preschool.

A fourth day of recording (about 5.5 hours) occurred a week later. On this day, the voice dosimeter was placed on the child early in the morning before being taken to the acoustic laboratories. During this final day, the physical environment was familiar to the child and his time was primarily spent speaking with adults, all of whom were at least somewhat familiar to the child. The structured acoustic environment measurement and tasks (described above) occurred towards the end of this period (about 11 am). Vocalizations from these tasks were not analyzed as part of the voice dosimeter study, but were separately analyzed to provide a baseline structured F_0 measure for comparison with the F_0 measures from the long-term, unstructured voice dosimeter data. The voice dosimeter was removed immediately after the structured tasks were administered, and no further dosimetry data was collected.

Observations by adults familiar with the child (i.e., parents, relatives, teachers, and family friends) reported no noticeable changes in his activities because of the device. Overall, the child quickly adapted to the device backpack. He appeared to play and interact with other children and adults as he normally would. Perhaps because the instrument was so familiar to him given his father's extended use of it, he also appeared to pay little attention to the dosimeter. Only once during the study (Day 3) did the child appear to be significantly conscious of the device and played with it. When this occurred, the data collection was ended for that day and the 2 hours before were excluded, limiting that day's usable data to 7 hours. Fundamental frequency during all other times was captured for analysis.

3. Analysis and Results

Fundamental frequency data from the structured vocal tasks performed in the controlled acoustic environment were obtained from analysis of both the voice dosimeter data and the simultaneously recorded head-mounted microphone signal. Fundamental frequency data from the unstructured activities were obtained only from analysis of the voice dosimeter data.

3.1. Results from Structured Vocal Tasks in a Controlled Acoustic Environment

For the entire in-booth recording protocol, the child wore a head-mounted microphone and the voice dosimeter. From the microphone signal, all vocalizations made by the child were parsed from the wave file for individual analysis. A new file (Figure 3) was created for analysis using the individual segmented files. Fundamental frequency was extracted from the microphone signal on all recorded segments mentioned above and on one-second segments of /a/ from *ahhh* and *mop* (after Baker et al.). Two methods were used: [1] TF32 software package (<http://userpages.chorus.net/cspeech>, which uses zero-crossing of an LPC inverse filtered time signal); and [2] NCVS analysis computer scripts (Figure 3, depicts graphic interface with the scripts), written in MATLAB with F_0 extraction (using subharmonic to harmonic ratio) based on Sun [27].

When analysis was conducted on the tasks as a whole and not just selected portions, both TF32 and the NCVS MATLAB programs first judged voicing segments of the recording to determine on which to do the analysis, with TF32 using default settings (based on signal amplitude). The MATLAB program judged voicing by an amplitude threshold of the part of the signal with energy within the range common to vocal fundamental frequency. Table 2 contains descriptive statistical metrics of extracted F_0 from the four structured vocal tasks as a whole and from segmented /a/ vowels. These metrics (e.g., mean, standard deviation, mode) were obtained from the combined extracted F_0 data of all three renditions of a particular task.

By comparing the F_0 extraction from the voice dosimeter and the head-mounted microphone, we can examine results from the three techniques for capturing F_0 . While all techniques extract F_0 , they do so using different methods. For example, although all of these techniques extract F_0 over a short window of time, none use the same size time window or the same overlap to calculate fundamental frequency: [1] The voice dosimeter uses a 30 ms window, an FFT algorithm and a peak picking extraction technique; [2] MATLAB scripts used 100 ms time windows sliding in 4 ms steps in time; and [3] TF32 uses an adaptive window size and step. The MATLAB and TF32 analysis used an overlapping window, where the dosimeter does not. Figure 4 illustrates how such distinctions might produce slightly different results. Here can be seen the extracted F_0 results of the second /ball/ rendition as obtained by the MATLAB scripts and by the NCVS dosimeter simultaneously. While both result in nearly the same F_0 contour, there are slight differences in the value of the extracted F_0 , as well as the number of extracted values for the same utterance.

Fundamental frequency distribution plots of the two microphone extraction methods for the entire concatenated set of recorded vocal tasks, as well as the dosimeter measurement over the entire structured recording time, are shown in Figure 5. Statistical metrics of the distributions are presented in Table 3. Although the F_0 distribution for each extraction technique is similar, there are two differences. First, there are inconsequential non-similar scales of the F_0 occurrence totals on the y-axis caused primarily by the difference in the analysis window which the three techniques use (discussed above). Second, each extraction method is different in its mathematical construct and frequency resolution, resulting in some slight variations in the overall distribution shapes. These variations may be explained by differences in extracted pitch during voicing onset and offset from inherent differences in the three extraction techniques' algorithm voicing detection and F_0 extraction methods. In addition, they are also likely caused

by the fact that the dosimeter extracted data also included any sporadic vocalizations made during the recording interval while such unelicited vocalizations were not used in the segmented microphone signal analysis. The decision was made to include these vocalizations in the dosimetry analysis of the structured tasks for comparative as well as practical reasons: [1] by including all vocalizations, the dosimeter study of the structured tasks paralleled the long-term dosimeter study, which also analyzed all of the child's vocalizations; and [2] as the voice dosimeter was designed to capture and tally all voice use, it was not practical to remove the non-task vocalizations that were extraneous to the protocol.

Despite these two differences, the two microphone signal extracted plots and the dosimeter produce very similar statistical metrics (see Table 3). As an additional check of the dosimeter extraction, the full microphone signal (including unelicited vocalizations) was also analyzed by TF32 and listed in Table 3. Further, given the differences in the extraction methods, the slight differences in the results are small and are not unexpected.

Table 2 indicates that the mean, median and mode were similar for each of the individual, structured elicited tasks. In fact, the child produced the *ahhh* and *mop* tasks with nearly identical means, medians, and modes for both microphone-based extraction techniques. When a 1-second middle section of the vowels were extracted from the *ahhh* and *mop* tasks, the results were little different. Interestingly, the child's F_0 for the *ball* and *count* tasks had lower means, medians and modes, as well as greater standard deviations than the other two tasks. These differences were, of course, caused by differences in the production of the tasks. For example, during the 3rd *count* task the child strongly emphasized '9' and '10', which included raising the pitch (Figure 3), a pattern absent from the first two renditions. In contrast, the median and mode (as opposed to mean and standard deviation) were more resistant to unique events and outliers and, thus, more likely preserved the general trend of the F_0 .

The child's F_0 for the tasks was within the range of similar-aged children found in Baker et al., with the child in the current study slightly above the average overall. The relative change from one task to the next (*ahhh*, *mop*, and *ball*) for both the child and the Baker et al. children were similar. Only the *count* task differed: the mean for this task was lower than the other tasks for the child, an opposite effect than what was found by Baker et al. However, the mean values for the *count* task in the two studies were nearly identical (242 Hz vs. 248 Hz).

The mean and mode of the combined extracted F_0 from structured tasks from the two methods of analyzing the microphone data (i.e., TF32 software and MATLAB analysis scripts) and from the dosimeter data were comparable, with the mean across the analysis techniques ranging from 257.4–259.9 Hz and the mode ranging from 258.4–259.9 Hz. In addition, these results (Table 3) were all comparable to the results from the same tasks as previously reported (mean F_0 range of about 217–262 Hz for 5-year-old children [12]), as well as the body of literature for this age (Table 1). Finally, as mentioned above, the distribution from three of the four combined structured tasks appeared to be nearly a normal distribution, so that mean and standard deviation would be proper descriptor statistics.

3.2. Results from Extended Dosimeter Study of Unstructured Vocalizations

The distribution and statistics calculated from the long-term unstructured vocalizations (34 hours) have quite different results (Table 4). Figure 6 illustrates both the mean and the median. Where the mean F_0 from the elicited tasks is approximately 260 Hz, the mean F_0 from the long-term, spontaneous speech is about 376 Hz. Another distinguishing trait is the large standard deviation (s.d. 115 Hz, compared to 19–28 Hz). Further, the median is also very high in the long-term vocalizations, approximately 355 Hz. Looking at Figure 6 we see that, while 355 Hz is the median F_0 value, the values reach ranges above 600 Hz on the high end and not less than 200 Hz on the low end. This means that the F_0 distribution is not a normal distribution,

heavily skewed to the right, and statistical mean and standard deviations would not be useful metrics.

Other statistical measures provide further detail into the distribution. First, by calculating mode, we learn that the most frequently occurring F_0 in the child was nearly 290 Hz, a value closer to the literature's reported F_0 values for this age during structured clinical settings. Further, the F_0 values that occurred at least as frequently as 75 percent or more of the mode provides us with the most frequent 75% range of F_0 values (between 240 and 350 Hz). Expanding this range out further, the range of the child's F_0 values which occurred at least 50% of the mode was approximately 240–480 Hz. It is to be noted that the upper part of this commonly occurring range was well above anything the child produced in the controlled recording environment, where there were few instances of frequencies above even 300 Hz. However, because of these higher fundamental frequencies, the child's mean F_0 was skewed upwards. Thus, it was significantly higher than what would be expected from not only his unstructured mode (290 Hz), but also his mean F_0 from the four structured tasks (~260 Hz).

4. Discussion

By nature, most child voice studies obtain speech samples in an unnatural environment (e.g., structured or unstructured elicitations, observations of play, etc.). Further, they have traditionally observed primarily speech F_0 . The value of the current study is that, even if changes in behavior and language were made by the child because of the voice dosimeter, these changes would likely be for a short time after attachment and only a small amount in comparison to the overall recording time. Thus, the impact of such initial changes would be reduced by the longevity of the study. Further, this study also uniquely recognizes the significance of non-speech vocalizations (i.e., squeals, screams, laughs, etc) within child communication and interaction by examining the non-speech vocalizations in tandem with speech-only vocalizations.

It is somewhat difficult to place the current study data within the context of the body of previous research (from Table 1). First, as discussed in the introduction, there are inconsistencies (and many ambiguities) in the ways that many of these studies report their results. Kent [28] suggests that there are limitations in using the body of child F_0 literature to make general predictions because the studies (up to that point) had used assorted study protocols and analyses. The contention was given that this diversity made it difficult to ascertain whether results demonstrated actual trends in specific age groups, or whether they illustrated the inter-subject variability in a group of subjects known for F_0 inconsistency. Since this study nearly 30 years ago, child F_0 studies have consistently improved. Nevertheless, they still use a variety of study designs, and few studies attempt to limit the inter-subject variability Kent described (e.g. [12]). Despite the potential ambiguities in the body of child F_0 studies, the child's F_0 during structured, task-based activities in a clinical environment, as measured by the dosimeter and the microphone, was within the boundaries of previous research.

A second reason for the difficulty in comparing the current data to the literature is that no previous studies have been found of long-term unstructured speech outside a clinical environment for a native English-speaking, 5-year-old male child. However, studies of younger children provide clues to an important pattern: mean F_0 in unstructured conversations appears to be notably higher than mean F_0 in structured conversations. Robb and Saxman [13] measured mean F_0 of 2 year olds ($N=3$) from free speech in a clinical environment to be 314 Hz (s.d. 60, range 268–364). Later, Hall et al. [7] found a similar mean for free speech F_0 in a clinical environment in children just older (mean age = 3 years, 10 months; $N=10$), 310 Hz with no range provided. On the other hand, Assman and Katz [9] elicited structured speech in 3 year olds ($N=10$) and found a much lower mean of 232 Hz (s.d. 19, range 209–271). The difference

in free speech F_0 and elicited speech F_0 is stark in these three examples. The very low mean F_0 produced by the children in Assman and Katz might be indicative of the extremely structured nature of the tasks: the children were asked to repeat monophthongal vowels as given to them from recorded adult productions. Thus, it is possible that the much lower mean F_0 was the result of the children attempting to mimic the pitches they heard.

By examining one child in a variety of environments, the current study adds valuable evidence by eliminating the weakness of inter-subject variability which Kent [28] describes. In doing so, this study demonstrates a clear pattern: the mode F_0 produced by a 5 year old in an unstructured, spontaneous (i.e., completely un-elicited) environment is considerably higher than the mean F_0 the same child produces in a structured, clinical evaluative environment. Although this difference likely stems primarily from the inclusion of an abundance of non-speech vocalizations in the long-term dosimetry study, it is also possible that the child lowered his F_0 to match the elicitor in the structured, clinical environment ([9]). Nevertheless, the current study does not provide enough data to explore this possibility.

By capturing not only structured responses in a clinical environment but also the entirety of unstructured vocalizations, this study also provides valuable data concerning the entirety of spontaneous child vocalizations. For example, the current subject's mode F_0 (i.e., his most commonly used F_0) in unstructured activities was comparable to the mean F_0 from Robb and Saxman [13] and Hall et al. [7]. However, his mean F_0 was considerably higher, skewed by some extremely high vocalizations (including squeals, screams, whining and other vocalizations) which, if they were elicited within a clinical environment, would not likely be produced to the same extent as in spontaneous, unstructured vocalizations. Instead, the child's extremely high vocalizations can be compared to the research of Nwokah et al. [15], whose examination of child laughter (described above) may justify the frequently occurring high F_0 produced by the current subject.

In addition to the higher F_0 produced in unstructured speech, the variation in F_0 range is also crucial to note. With a mean of 314 Hz, the children in Robb and Saxman's study produced mean frequencies with nearly a 100-Hz range. When the researchers presented the data grouped by age in 3-month intervals along with the variation coefficient, the wide range in F_0 produced by the children was demonstrated (269–537 Hz). This pattern is also seen even in studies of structured speech in older children. For example, Robb and Smith [14] recorded frequencies in 4+-year-old boys as high as 323 Hz; Hasek and Singh [8] recorded frequencies as high as 313 Hz in 5-year-old boys and 332 in 6-year-old boys; and Weinberg and Zlatin [11] recorded frequencies as high as 293 Hz in 5+-year-old boys. These maximum F_0 values are not only comparable to those produced by the child in his structured tasks, but also fall within his most frequently occurring F_0 values (75% of mode) in the long-term, unstructured, dosimeter study (Figure 5).

Overall, these results point to a potential significant weakness in studies that rely on mean and standard deviation (as well as range when reported as range of means and range of standard deviations) to accurately capture unstructured F_0 in children. From the current study, the F_0 captured during the elicited speech from three of the four elicited tasks had a nearly normal distribution with a similar mean, median and mode. In such a situation, mean and standard deviation are likely valid metrics which would accurately describe the child's vocalizations. However, the mean would not be the best statistic for any non-normal results, such as the fourth elicited task in the current study (i.e., *count*), where the mean and mode are quite different (241.7 Hz vs. 221.6 Hz). Although the mode is not provided, it is possible that a similar inconsistency might be present in Baker et al. given its similarity in mean to the current study. Further, it is possible that a similar inconsistency might be seen in the raw data from other studies, such as Hasek and Singh [8], whose reported range of F_0 for 5-year-old boys was 186–

313 Hz. It is likely that, with a range so large, the mean, median, and mode would have been dissimilar.

Building on this point, this study also highlights potential problems in the use of mean and standard deviation when examining F_0 collected from long-term unsolicited speech. With nearly a 100 Hz difference between the long-term median and mode F_0 , both of these metrics would provide an incomplete picture of the child's true vocalization patterns. Further, although maximum and minimum F_0 is often quite useful to capture a picture of a subject's voice use, it would not be appropriate with long-term histogram distributions of a child's voice such as found in the current study because of the likelihood that a single extreme outlier would occur over tens of thousands of F_0 productions. In such a situation, a statistical metric such as 75% or 50% of mode might provide a better insight into the range of F_0 values used.

In addition, this study points to the importance of reporting the results clearly, consistently and completely. For example, Awan and Mueller [6], provided a comprehensive picture of the children which they examined by providing not only the mean and standard deviation of the speaking F_0 (Hz), but also: (1) maximum and minimum speaking F_0 ; (2) standard deviations of the means; and (3) mean of the standard deviations. The researchers also provided each of these same statistics for semitones, allowing for comparison to other previous studies which give only semitones or a combination of semitones and F_0 (e.g., Eguchi & Hirsh [29] and Weinberg & Bennett [30]). Further, Nwukah et al. [15] provide a valuable pattern in their research by not only providing a wide variety of data (i.e., mean F_0 , maximum and minimum F_0 , and F_0 span, as well as the standard deviations for each), but also providing a table which defines how each statistical metric was calculated.

5. Conclusions

Acoustical measures of voice and speech such as fundamental frequency (F_0) are useful for assessing a child's voice. The current study explores the following research questions to compare child F_0 production. First, does a child produce similar F_0 patterns during structured vocal tasks in a clinical evaluative environment (i.e., elicited responses and few non-speech vocalizations), as during long-term unstructured activities (i.e., both speech and non-speech vocalizations)? Second, does the F_0 during these activities follow a normal distribution, thereby allowing usual basic statistical metrics (i.e., mean and standard deviation) to be properly used?

The current study affirms previous measures of a 5-year-old child's speech F_0 during elicited structured speaking tasks [12], reaffirming the suggestion that more consistent F_0 elicitation practices be adopted to eliminate the inconsistency caused by methodology and analysis variability (Baker et. al, [12]; Kent, [28]). Nevertheless, this study also suggests that the common use of mean F_0 value as a statistical metric may not always be appropriate in child F_0 studies. While three of the four structured tasks in the current study had nearly normal F_0 histograms, the fourth task (i.e., counting) was not as normal. With the mode quite different from the mean, a full understanding of how the child produced the task would not be captured by simply reporting mean F_0 .

Further, this study demonstrates that mean F_0 and standard deviation are likely inadequate indicators of a child's F_0 during studies of long-term, spontaneous vocalizations in a non-clinical setting. During such studies, it is important to pinpoint the child's most commonly used F_0 (or the most commonly used F_0 range), as well as the full range which the child uses (i.e., $F_{0\min}$, $F_{0\max}$). Further, it is necessary to report data in a way to represent the skewed distribution likely found in long-term vocal patterns. To simplify, mode and median are proposed as the statistic metrics of long-term F_0 . In addition, a metric which depicts the range

in each subject's actual F_0 (in addition to simply the range of means) should be presented (e.g., quartiles, percentage of mode, or mean of standard deviations).

Finally, the current study showed an unambiguous difference in vocalizations during structured, elicited tasks versus actual F_0 during long-term, unstructured vocalizations. This likelihood has important clinical implications. While structured, consistent F_0 elicitation may be used to accurately track therapy progress, they may not be at all representative of a child's everyday, unelicited, unstructured voice use. Thus, in terms of observing pediatric voice use and vocal health, methods for more accurately measuring typical speech F_0 might need improvement. New devices, such as the National Center for Voice and Speech voice dosimeter, now make such studies which would build on the results of the current study logistically more feasible than in the past.

Although this study does not give a complete explanation, these differences in F_0 are caused at least in part by the degree of non-speech vocalizations which are common attribute of a child's daily vocalizations but which are likely not fully captured in a structured, clinical setting. It is also possible that a child may alter his/her F_0 in an attempt to match the communication environment in which they find themselves (e.g., Assman and Katz [9]). With both of these possibilities in mind, it is crucial to step beyond the traditional child F_0 studies which have examined the effect of task-type, age, gender and race to explore specific conditions in both structured and unstructured environments. First, how is F_0 affected in structured clinical settings by the testing environment, the familiarity of the child with the investigator, and by the age/gender of the investigator? Second, how is F_0 affected in unstructured environments by such conditions as: [1] the vocal environment (e.g., home, classroom, or playground); [2] the familiarity of the child with the communication partner(s) (e.g., family member, relative, friend, or new associate); [3] the level of perceived authority (e.g., parent, teacher, adult or older child); and [4] the number and age(s) of the partner(s).

It is also important to further explore the pattern of high F_0 in children, in particular to capture to what extent they are part of a child's spontaneous, unstructured vocalizations. Also valuable would be an examination of intensity, specifically parsing out how a child's intensity might vary with speech vocalizations compared with non-speech vocalizations (e.g., squeals, screams, laughs, animal or truck noises, etc). It would also be important to track how a subject's long-term vocalizations change over multiple years, for example how the use of non-speech vocalizations might change as the child ages. Finally, the results found in this study were from a single subject and should be confirmed by additional long-term studies with more subjects from both genders and of different ages. In this way, we can further delineate what child F_0 data demonstrate actual trends in specific age groups, and what stem from inter-subject variability.

Acknowledgments

Funding for this work was provided in part by the National Institute on Deafness and Other Communication Disorders, grant number 1R01 DC04224, P.I. Ingo R. Titze. The author would like to thank the research team (both past and present) at the National Center for Voice and Speech with many supporting roles (Dosimeter Team: Jan Svec, Peter Popolo, Andrew Starr, Albert Worley; General contributions: Jennifer Spielman, Angela Halpern). Thank you to Laura M. Hunter for literature search and technical review of the document.

Support: NIDCD R01-DC004224

Reference List

1. Wertzner HF, Schreiber S, Amaro L. Analysis of fundamental frequency, jitter, shimmer and vocal intensity in children with phonological disorders. *Rev Bras Otorrinolaringol (Engl Ed)* 2005;71:582–588.

2. Munson B, Bjorum EM, Windsor J. Acoustic and perceptual correlates of stress in nonwords produced by children with suspected developmental apraxia of speech and children with phonological disorder. *J Speech Lang Hear Res* 2003;46:189–202. [PubMed: 12647898]
3. Scheiner E, Hammerschmidt K, Jurgens U, Zwirner P. Acoustic analyses of developmental changes and emotional expression in the preverbal vocalizations of infants. *J Voice* 2002;16:509–529. [PubMed: 12512639]
4. Crary MA, Tallman VL. Production of linguistic prosody by normal and speech-disordered children. *J Commun Disord* 1993;26:245–262. [PubMed: 8126262]
5. Bohme G, Stuchlik G. Voice profiles and standard voice profile of untrained children. *J Voice* 1995;9:304–307. [PubMed: 8541974]
6. Awan SN, Mueller PB. Speaking fundamental frequency characteristics of white, African American, and Hispanic kindergartners. *J Speech Hear Res* 1996;39:573–577. [PubMed: 8783135]
7. Hall KD, Yairi E. Fundamental frequency, jitter, and shimmer in preschoolers who stutter. *Journal of Speech and Hearing Research* 1992;35:1002–1008. [PubMed: 1447911]
8. Hasek CS, Singh S, Murry T. Acoustic attributes of preadolescent voices. *Jounral of the Acoustical Society of America* 1980;68:1262–1265.
9. Assmann PF, Katz WF. Time-varying spectral change in the vowels of children and adults. *J Acoust Soc Am* 2000;108:1856–1866. [PubMed: 11051512]
10. Lee S, Potamianos A, Narayanan S. Acoustics of children's speech: developmental changes of temporal and spectral parameters. *J Acoust Soc Am* 1999;105:1455–1468. [PubMed: 10089598]
11. Weinberg B, Zlatin M. Speaking fundamental frequency characteristics of five- and six-year-old children with mongolism. *J Speech Hear Res* 1970;13:418–425. [PubMed: 4246267]
12. Baker S, Weinrich B, Bevington M, Schroth K, Schroeder E. The effect of task type on fundamental frequency in children. *Int J Pediatr Otorhinolaryngol* 2008;72:885–889. [PubMed: 18395803]
13. Robb MP, Saxman JH. Developmental trends in vocal fundamental frequency of young children. *J Speech Hear Res* 1985;28:421–427. [PubMed: 2981049]
14. Robb MP, Smith AB. Fundamental frequency onset and offset behavior: a comparative study of children and adults. *Journal of Speech and Hearing Research* 2002;45:446–456.
15. Nwokah EE, Davies P, Islam A, Hsu HC, Fogel A. Vocal affect in three-year-olds: a quantitative acoustic analysis of child laughter. *J Acoust Soc Am* 1993;94:3076–3090. [PubMed: 8300948]
16. Baeck HE, de Souza MN. Longitudinal study of the fundamental frequency of hunger cries along the first 6 months of healthy babies. *J Voice* 2007;21:551–559. [PubMed: 16730945]
17. Rothganger H. Analysis of the sounds of the child in the first year of age and a comparison to the language. *Early Hum Dev* 2003;75:55–69. [PubMed: 14652159]
18. Varallyay G Jr, Benyo Z, Illenyi A, Farkas Z, Kovacs L. Acoustic analysis of the infant cry: classical and new methods. *Conf Proc IEEE Eng Med Biol Soc* 2004;1:313–316. [PubMed: 17271673]
19. Akif KM, Okur E, Yildirim I, Guzelsoy S. The prevalence of vocal fold nodules in school age children. *Int J Pediatr Otorhinolaryngol* 2004;68:409–412. [PubMed: 15013605]
20. Shah RK, Woodnorth GH, Glynn A, Nuss RC. Pediatric vocal nodules: correlation with perceptual voice analysis. *Int J Pediatr Otorhinolaryngol* 2005;69:903–909. [PubMed: 15911007]
21. Svec JG, Popolo PS, Titze IR. Measurement of vocal doses in speech: experimental procedure and signal processing. *Log Phon Vocol* 2003;28:181–192.
22. Svec, JG.; Hunter, EJ.; Popolo, PS.; Rogge-Miller, K.; Titze, IR. NCVS Memo No 02. The Calibration and Setup of the NCVS Dosimeter. NCVS Online Technical Memo; April 2004; 2004.
23. Popolo PS, Svec JG, Titze IR. Adaptation of a Pocket PC for use as a wearable voice dosimeter. *Journal of Speech Language and Hearing Research* 2005;48:780–791.
24. Titze IR, Hunter EJ, Svec JG. Voicing and silence periods in daily and weekly vocalizations of teachers. *J Acoust Soc Am* 2007;121:469–478. [PubMed: 17297801]
25. Carroll T, Nix J, Hunter E, Emerich K, Titze I, Abaza M. Objective measurement of vocal fatigue in classical singers: A vocal dosimetry pilot study. *Otolaryngol Head Neck Surg* 2006;135:595–602. [PubMed: 17011424]
26. Titze IR, Svec JG, Popolo PS. Vocal dose measures: quantifying accumulated vibration exposure in vocal fold tissues. *Journal of Speech, Language, & Hearing Research* 2003;46:919–932.

27. Sun, X. Pitch determination and voice quality analysis using subharmonic-to-harmonic ratio. IPASSP2002; Orlando, Florida. 2002.
28. Kent R. Anatomical and neuromuscular maturation of the speech mechanism: evidence from acoustic studies. J Speech Hear Res 1976;19:421–447. [PubMed: 979206]
29. Eguchi S, Hirsh IJ. Development of speech sounds in children. Acta Otolaryngol Suppl 1969;257:1–51. [PubMed: 5265922]
30. Weinberg B, Bennett S. Speaker sex recognition of 5- and 6-year-old children's voices. Journal of the Acoustical Society of America 1971;50:1210–1213.

Figure 1. The NCVS voice dosimeter worn by a male adult (left) and sample output (right) of loudness, frequency (spectral centroid and fundamental frequency), and voicing detection in 30 msec intervals.

Figure 2.
The child with the NCVS voice dosimeter accelerometer attached.

Figure 3.

Using custom computer scripts, the three renditions of each vocal task (shown after trimming and grouping each task) were analyzed and the F_0 was extracted. The top panel illustrates the pressure waveform, the middle panel illustrates the relative dB, the F_0 contour is shown in the lower panel.

Figure 4.

The F₀ contour extracted from the word 'ball'. The squares indicate the F₀ as found by the dosimeter, while the connected points were found from the microphone data.

Figure 5.
F₀ distribution of the structured vocal tasks as measured by three different methods.

Figure 6.
Fundamental Frequency distribution of a 5-year-old boy over 34 hours of observation.

Table 1

Summary of F_0 studies, arranged by age of subjects. Limitations: Native U.S.-English speaking; $\geq 2;0$, $\leq 8;0$ (years; months). Studies with no clear delineation of F_0 and/or age were excluded. Studies without listing age in year and months are listed below as year only. Note that standard deviation (**SD**) and range of F_0 are usually the range of the mean across subjects.

Study	Voice/Speech Sample Type	Number Subjects	Age/Gender	Mean F_0 (Hz)	SD (Hz)	Range F_0 (Hz)
Robb & Saxman (1985)	Free speech using familiar books, toys; mean F_0	3	2;1/N/A	314	60 ^{†††††}	268–364
	Free speech using familiar books, toys; F_0 onset	3	2;1/N/A	325	60 ^{†††††}	272–378
	Free speech using familiar books, toys; F_0 across vowel segments	3	2;1/N/A	314	60 ^{†††††}	269–364
Robb et al. (1989)	Monosyllable and bisyllable from spontaneous interaction with mother and investigator	3	2-2;2/both	374	85 ^{†††††}	297–453
Eguchi & Hirsh (1969)	6 vowels extracted from sentence	5	3/both	298	31	---
Nwokaah et al. (1993)	Acoustic analysis of seven types of laughter.					
	Dull Comment			395	92	270–480
	Exclamatory Comment			416	69	380–510
	Chuckle			439	82	340–600
	Basic Rhythmical	4	3/both	470	73	360–650
	Variable Rhythmical			411	75	340–540
Assmann & Katz (2000)	Classical Rhythmical			494	42	390–560
	Squeal			1907	602	1300–2100
Hall and Yairi (1992)	12 monophthongal vowels of American English in /hVd/ context	10	3/both	232	19 22 ^{†††††}	209–271
Eguchi & Hirsh (1969)	Vowels from free speech	10	3;10/males	310	33	---
	From 6 vowels extracted from five repetitions of a sentence	5	4/both	286	21	---

Study	Voice/Speech Sample Type	Number Subjects	Age/Gender	Mean F0 (Hz)	SD (Hz)	Range F0 (Hz)
†††† Perry et al. (2001)	7 nondiphthongal vowels of American English in the phrase, "Say /hVd/ again."	10	4;3/males	242	4	---
		10	4;3/males	236	6	
† Robb & Smith (2002)	Vocal onset, first 10 cycles	10	4;5/both	292	39	285 – 323
	Vocal offset, first 10 cycles	10	4;5/both	284	40	276 – 290
Weinberg & Bennett (1971)	Spontaneous responses to uniform pictures and questions (30-sec)	15	5/males	252	45	217 – 293
		18	5/females	258	46	212 – 295
Eguchi & Hirsh (1969)	From 6 vowels extracted from five repetitions of a sentence	5	5/both	289	46	---
Assman & Katz (2000)	12 monophthongal vowels of American English in /hVd/ context	10	5/both	259.9	13	240 – 280
†††† Baker et al. (2008)	4 tasks, repeated 3 times	15	5/both	240	---	215 – 265
Hasek & Singh (1980)	Sustained /a/	15	5/males	248	28	186 – 313
		15	5/females	258	29	202 – 306
Lee et al. (1999)	10 monophthongal vowels, 5 sentences	19	5/males	267	9	260 – 287
		13	5/females	274	12	258 – 300
Weinberg and Zlatin (1970)	Speech elicited from uniform questions, pictures, and puzzles	15	5;7/males	252	3	217 – 293
		18	5;5/females	248	3	212 – 295
†† Awan & Mueller (1996)	Speech elicited from uniform picture-description task	18	5;6/males	241	18	205 – 274
		17	5;7/females	231	15	208 – 262
Awan & Mueller (1996)	Speech elicited from uniform picture-description task	15	5;7/males	240	16	212 – 263

Study	Voice/Speech Sample Type	Number Subjects	Age/Gender	Mean F0 (Hz)	SD (Hz)	Range F0 (Hz)
<i>††††</i> Baker et al. (2008)	4 tasks, repeated 3 times	20	5;7/females	243	22	195 – 291
<i>††††</i> Eguchi & Hirsh (1969)	From 6 vowels extracted from five repetitions of a sentence	5	6/	271	28	
Weinberg & Bennett (1971)	Spontaneous responses to uniform pictures and questions (30-sec)	14 19	6/males 6/females	247 247	3 3	204 – 274 218 – 274
Hasek & Singh (1980)	Sustained /a/	15 15	6/males 6/females	263 254	32 23	228 – 332 210 – 313
<i>††</i> Wheat & Hudson (1988)	Prompted spontaneous speech	50 50	6/males 6/females	220 211		135 – 299 138– 298
Sorenson (1989)	Speech elicited from 3 tasks: Oral reading, Spontaneous speech about specific picture, Sustained vowel productions	3 3	6/males 6/females	269 308		253 – 285 294 – 325
Lee et al. (1999)	10 monophthongal vowels, 5 sentences	11 16	6/males 6/females	267 264	11 11	256 – 290 248 – 282
<i>†††</i> Vuorenkoski et al. (1978)	Repeat sentence with long vowels; glide voice low as possible; glide voice high as possible	38 22	6;4/males 6;5/females	261, 234 256, 230		
Weinberg & Zlatin (1970)	Spontaneous speech elicited uniform questions, pictures, and puzzles	14 19	6;6/males 6;4/females	247 247	3 33	204 – 274 218 – 274
<i>††††</i> Eguchi & Hirsh (1969)	From 6 vowels extracted from five repetitions of a sentence	5	7/	263	39	

Study	Voice/Speech Sample Type	Number Subjects	Age/Gender	Mean F0 (Hz)	SD (Hz)	Range F0 (Hz)
Sorenson (1989)	Speech elicited from 3 tasks: Oral reading, Spontaneous speech about specific picture, Sustained vowel productions	3 3	7/males 7/females	288 279	27 35	273 – 310 258 – 291
Fairbanks (1950)	Out loud reading of 52 word passage	15 15	6;10-7;2/males 6;10-7;2/females	294 277	22 26	
Assman & Katz (2000)	12 monophthongal vowels of American English in /hVd/ context	10	7/both	248.3	8	235 – 257
†††† Baker et al. (2008)	4 tasks, repeated 3 times	18	7/both	236	---	208 – 257
Hasek & Singh (1980)	Sustained /a/	15 15	7/males 7/females	234 262	18 36	199 – 264 195 – 303
†††† Lee et al. (1999)	10 monophthongal vowels, 5 sentences	11 24	7/males 7/females	266 277	9 9	253 – 281 265 – 298

† Possibly natural frequency before adjusting because of linguistic, cultural, or environmental pressures.

†† Paper specified that the children were African-American.

††† Paper specified the subjects' Speaking Fundamental Frequency (SFF) and Lowest Frequency (LF), included in table respectively.

†††† Data was read off graphs presented in manuscripts and should be taken only as approximate.

††††† Average of standard deviations.

F₀ statistics for the controlled acoustic environment vocal tasks as extracted from the microphone. Results show a combination of each of the individual tasks (Full) and /a/ vowel only (Vowel).

Utterances	Full (MATLAB)				Full (TF32)				Vowel (TF32)	
	ahhh	mop	ball	count	ahhh	mop	ball	count	ahhh	mop
mean	263.4	260.1	252.0	241.7	260.3	260.4	243.7	254.6	263.2	260.9
std dev	9.9	11.1	25.0	25.0	12.7	11.4	31.3	25.2	6.0	8.6
median	260.5	260.5	249.5	237.6	260.3	260.3	242.2	252.3	261.8	260.4
mode	259.1	260.5	249.5	221.6	260.6	260.6	234.2	248.7	256.9	259.9

Table 3

F₀ statistics across all of the vocal tasks obtained by the three different methods.

All Tasks, All Renditions	MATLAB (tasks only)	TF32 (tasks only)	TF32 (all)	Dosimeter (all)
<i>mean</i>	257.6	257.4	256.5	259.9
<i>std dev</i>	18.7	28.2	21.8	25.0
<i>median</i>	259.1	259.1	258.9	258.4
<i>mode</i>	259.1	259.9	258.7	258.4

Table 4

The statistics for the long-term F_0 distribution (34.1 hours of recording) for the child. Over that time, the child voiced 21.1% of the time.

Statistic	
mean:	376.5
variance:	13222.0
<i>standard deviation:</i>	115.0
median:	355.3
quartile 1:	290.7
quartile 2:	441.4
mode:	290.7
skew:	1.1
bowley_skew:	0.1
kurt:	4.1