Published in final edited form as:

Lang Speech Hear Serv Sch. 2012 July ; 43(3): 253–263. doi:10.1044/0161-1461(2011/11-0002).

Impact of four non-clinical speaking environments on a child's fundamental frequency and voice level: A preliminary case study

Eric J. Hunter^{1,2,3}, Angela E. Halpern⁴, and Jennifer L. Spielman⁴

¹National Center for Voice and Speech, The University of Utah, Salt Lake City, UT

²Department of Bioengineering, The University of Utah, Salt Lake City, UT

³Division of Otolaryngology – Head and Neck Surgery, University of Utah School of Medicine, The University of Utah, Salt Lake City, UT

⁴National Center for Voice and Speech-Denver, Affiliate of the University of Colorado, Boulder, CO

Abstract

Purpose—The aim of the study was to investigate how a child's fundamental frequency (F_0) and estimated voice level (dB SPL) change in distinct speaking environments.

Methods—A child (5 years, 7 months) wore a National Center for Voice and Speech voice dosimeter for four days. The two parameters used were F_0 and dB SPL. During analysis, the F_0 and dB SPL data were segmented to represent four typical speaking environments of school-age children: [1] free-play (2.5 hours); [2] preschool (3 hours); [3] home (10.7 hours); and [4] adult (5.6 hours). Unique to this study, the child's voice data were presented as Voice Use Profiles.

Results—The child's F_0 and dB SPL patterns within an adult environment were similar to that found in the literature, but showed much greater variation in the free-play environment. The preschool environment elicited a lower modal F_0 than did the home, but a higher median and mean F_0 , as well as a somewhat elevated mean dB SPL.

Conclusions—The child produced significantly different F_0 and dB SPL patterns across four different speaking environments. If future studies substantiate this pattern, clinicians and researchers must be aware of this difference when working with children.

Keywords

fundamental frequency; voice level; child; voice use

INTRODUCTION

Studies of children's voices have traditionally focused on the analysis of a few acoustic measures, rather than the wide range of measures typically studied in adult voices. One such measure is fundamental frequency, or F_0 (e.g., Hall & Yairi, 1992; Awan & Mueller, 1996). Although F_0 does not provide a complete picture of vocal use, it does offer important information about a child's typical speaking pitch, vocal variability, and vocal range. Used clinically, a deviation in typical F_0 parameters may reflect underlying problems affecting the voice and provide insight into the potential source(s) of a specific voice disorder.

Nevertheless, the use of F_0 measures with children is not without problems. First, children's F_0 values reported in the literature vary considerably (see Hunter, 2009, for a comprehensive synopsis of F_0 studies of native U.S. English speaking children 2–7 yrs). The reasons for these variations are unclear, but they may partially reflect the type of task the child is asked to perform (Baker, Weinrich, Bevington, Schroth, & Schroeder, 2008). Whatever the reason, such wide variations make it difficult to place an individual child's voice use in the continuum between normal and abnormal.

Second, it is possible that F_0 values collected in clinical and/or laboratory environments may not reflect what a child actually produces in his/her daily vocalizations. Hunter (2009) reported a case study in which a young boy produced F_0 values in a structured environment that fell within the range of averages reported in the literature but became dramatically different when he vocalized outside of the research setting. If these data can be substantiated with more children, it may impact both clinical practice and study methodology. For example, a recent study of school age children compared laryngeoscopic examination results to a sustained voice sample (at least 1 second in duration) in a clinical setting. While 30% of the children had abnormal findings on laryngoscopic examinations (Akif, Okur, Yildirim, & Güzelsoy, 2004), no relationship was detected between their limited observations of voice use and these abnormalities (i.e., lesions, immature nodules, nodules, and polyps). One possible explanation for this lack of relationship is that the acoustic parameters in the before mentioned Akif et al. (2004) study were recorded in a clinical setting, which Hunter (2009) suggests may provide little indication of how the children actually used their voices.

Finally, the use of F_0 alone does not provide a complete picture of a child's typical vocal patterns. Consideration should also be given to the child's vocal level (dB SPL), which may be particularly important in school-age children who use a wide range of loudness in play versus classroom settings. Few studies appear to examine dB SPL either alone or in tandem with F_0 in children (e.g., Wuyts, Heylen, Mertens, Du, Rooman, Van de Heyning, & De Bodt, 2003; Heylen, Wuyts, Mertens, De Bodt, Pattyn, Croux, & Van de Heyning, 1998). Hacki and Heitmüller (1999) compared age effects of children (4–12 y/o) on Voice Range Profiles (maximum dB and F_0 range), in which they found both habitual pitch and intensity for children but only for the clinical environment. Previous studies with adult subjects suggest that increased dB SPL may affect vocal pitch. For example, Gramming, Sundberg, Ternstrom, Leanderson, & Perkins (1988) found that, across 38 individuals who had different levels of professional voice training and varied in their vocal health, there was a mean F_0 increase of approximately $\frac{1}{2}$ -semitone per dB. This relationship was further demonstrated using long-term F_0 and dB SPL measures in school teachers, whose pitch and loudness rose in parallel during occupational voice use (Hunter & Titze, 2010). It would be important to ascertain to what degree this relationship exists in children.

Thus, it seems possible that while F_0 captured within a controlled clinical or research environment offers a simple way to monitor a child's voice use, it may not provide a complete picture of how school-age children actually use their voices within the different communication environments in which they find themselves on a daily basis. Children are exposed to a wide range of communication environments and partners during language development. It is possible that they learn to adapt their voices to the needs of each.

To explore this potential adaptation, the current study builds on Hunter (2009). The first author's child¹, a native English-speaking male (age = 5 years, 7 months), was the subject of the original study. The child passed a basic hearing screening by a school audiologist. He also was informally observed by two speech-language pathologists (SLPs) during typical conversational interactions as having no resonance disorders or voice disturbances. The two SLPs have more than 10 years of research experience with voice and speech disorders, and

are authors of the current study. They were unrelated to the first author or son, but very familiar with the child's speech habits. He was engaged in age-appropriate educational programs, namely thrice weekly preschool classes during the dosimeter period, and was scheduled to begin kindergarten two months after the recordings were taken.

The National Center for Voice and Speech (NCVS) voice dosimeter was used to capture the child's voice use. The dosimeter was originally developed to measure long-term voice use in adult subjects. Specific analysis and data collection techniques have been previously explained elsewhere (Švec, Hunter, Popolo, Rogge-Miller & Titze, 2004; Švec, Popolo & Titze, 2004; Popolo, Švec & Titze, 2005). This device measures vocal dose (Titze, Švec & Popolo, 2003) by an accelerometer attached with surgical glue to the sternal notch, and produces calculations of F_0 , voicing duration, and vibration amplitude (which can be converted to estimated dB SPL, Švec, Titze & Popolo, 2005). Since the transducer is an accelerometer, only skin vibration (primarily from the vocal folds) is transduced and sound waves are not. The dosimeter can record more than 24 h of voice data with the help of an external battery before needing to be recharged.

The dosimeter was placed in a small, unobtrusive backpack so that the child could have nearly full freedom of movement without hindrance. The child wore the device for a total of four days. The first three days were in the same week and entailed three major activities. First, there was 2.5 hours of free play with 10 other children at a birthday party; during this time, accompanying adults were present but participated very little in the children's activities. This environment included very active play on indoor equipment (e.g., large slides and swinging ropes) located in a padded room at a children's play gym. The second set of data was collected during 3 hours in a small in-home preschool. The class consisted of one adult and seven children, with periods of formal classroom instruction and structured playtime (both inside and outside, depending on the weather). Third, another 10.7 hours of data were collected in the home with the family, which consisted of a father (the primary investigator) and mother, as well as two younger female siblings (ages 2 years, 11 months; 1 year, 5 months). Captured over all three days, this consisted of structured family time as well as free play. On the fourth and final day (which occurred one week later), the child wore the device for 5.6 hours. On this day, the child was primarily in an office environment with his father and other adults who were at least somewhat familiar to the child. The child was not given the opportunity to play during this time. Although some data were collected on this fourth day during the commute on public transportation into the office, no data were collected during time spent with family members or other children.

Additional data from this study first described in Hunter (2009) are reported in the current study. Specifically, the current study explores the following research question: How do a child's F_0 and dB SPL differ across distinct speaking environments? Such data would be useful as a starting point in launching further studies into child language development and voice and speech use.

METHODS

As mentioned above, the current study uses data originally collected during a previously reported case study (Hunter 2009). The NCVS voice dosimeter was used to capture the voice use of a native English-speaking male child (age = 5 years, 7 months), as described

¹As 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 immediately asked to have it removed; the son wanted to keep it on and, thus, the data collected were later analyzed for the current case study.

above (for further details, see Hunter, 2009). Fundamental frequency and estimated dB SPL were captured for analysis as reviewed previously. Both F_0 and vocal estimated dB SPL were extracted from the four days of data collected by the dosimeter.

For the current study, four communication environments were identified and uniquely analyzed: [1] 5.6 hours of controlled activity with adults in a restricted environment (*adult*); [2] 2.5 hours of relatively unrestrained play with other children (*free-play*); [3] 3 hours of moderately restrained activity in a preschool (*preschool*); and [4] 10.7 hours at home with immediate family members (*home*). These four environments were delineated to replicate general speaking environments of young school-age children. *Home* hours were combined from different days, based on the child's general activities. More specific delineations were not attempted because the extreme variations in activity type would otherwise make conclusions difficult to reach.

For each of these four environments, both F_0 and estimated dB SPL were extracted from the four days of data collected by the dosimeter. The data were analyzed to determine if there were differences in average F_0 and dB SPL produced during these four communication environments. Data preparation and analysis were accomplished using custom MATLAB computer scripts. Standard statistical metrics were obtained (e.g., mean, median, mode, and kurtosis). In addition, voicing percentages (percent of voicing instances over a given recording time) were also calculated for the specific communication environments.

Fundamental frequency and dB SPL data from specific communication environments were analyzed and presented as histograms. Quartiles (used to demarcate data into groups each containing a quarter of the data) were calculated to indicate range of both F_0 and dB SPL. Quartiles were used because: [1] it is common in F_0 extraction to have occasional errant values (e.g., period doubling), which would skew minimum/maximum values; and [2] the recorded F_0 values did not have a normal distribution (for traditional statistical deviation calculations like mean and standard deviation, a normal distribution is assumed). Because dB SPL is calculated from the amplitude of the voicing signal, it is not subject to the same type or number of errant values as F_0 . In addition, dB SPL has a more normal distribution when compared to F_0 . However, to increase consistency, the quartiles were used in addition to normal statistical metrics for dB SPL. The first quartile is the median of the data between the minimum value and the full data median, or the first 25th percentile; the second quartile is the median, or the 50th percentile; and the third quartile is the median of the data between the median and the maximum data value, or the 75th percentile. The inter-quartile range is the difference between first and third quartiles, or the middle 50% of the data. In the current study, the inter-quartile range was used to quantify the range of F_0 and dB SPL used by the child.

RESULTS

First, F_0 results are described for each of the four communication environments. Next, the vocal level (estimated dB SPL) data are reported. Finally, the relation between vocal F_0 and the estimated dB SPL level is shown in terms of a Voice Use Profile.

Fundamental Frequency

The distribution histograms for F_0 in the four communication environments are shown in Figure 1. From these distributions, statistical descriptors are presented in Table 1. The highest F_0 values were found in *free-play* (Figure 1.d): median, 409.1 Hz; mean, 423.4 Hz; inter-quartile range, 172 Hz. While a mode value was calculated (366.1 Hz), the *free-play* distribution illustrates the child used a wide range of F_0 values with nearly equal incidence (300–450 Hz). In contrast, the child's F_0 in the *adult* setting had the lowest values: median,

312.2 Hz; mean, 334.4 Hz, mode, 258.4 Hz. Further, he also displayed a more restricted inter-quartile range in this environment (108 Hz).

The other two communication environments fell between these two extremes. In the *preschool* environment, the child's F_0 distribution had a median of 376.8 Hz, a mean of 396.4 Hz, and a mode of 279.9 Hz. In the *home* communication environment, the median was 355.3 Hz, the mean was 378.7 Hz, and the mode was 290.7 Hz.

Figure 1 illustrates the results of F_0 usage between the four communication environments using the primary statistics from Table 1 (mean, median, mode and inter-quartile range). The environments are ordered according to median and quartile scores, resulting in a sequential order showing a steady increase in F_0 parameters from environment to environment. The one exception to this is the modal F_0 in the *preschool* environment, which is lower than the *home* mode F_0 . The child used his lowest F_0 during the *adult* environment and his highest F_0 most frequently in the *free-play* environment.

Estimated Voice Level

Distribution histograms for estimated dB SPL at 30 cm in the four environments are shown in Figure 3 below. Statistical descriptors of the distributions are presented in Table 2. From Table 2, we see that the lowest dB SPL occurred in the *adult* environment (mean, 64.1 dB SPL; median, 63.0 dB SPL; mode, 63.0 dB SPL). During this time, the inter-quartile range was the most restricted (7 dB), as was the variance (32.2 dB SPL). The mean dB SPL was also quite low (65.6 dB SPL) in *free-play*. As with the F_0 , a mode was calculated (63 dB SPL), but the child used a wide range of dB SPL values between 57 and 70 dB SPL, with no prominent peak. The child had the largest inter-quartile range (10 dB) in *free-play* and by far the largest variance (46.1 dB SPL), depicting the broad distribution.

The child used, on average and most frequently, a higher dB SPL in *preschool* than in the other environments (mean, 67.0 dB SPL; median, 67.0 dB SPL; mode, 68.0 dB SPL), but with a smaller range of dB SPL than used in *free-play*. The child's dB SPL at *home* was lower by a small amount (mean, 66.2 dB SPL; median, 65.6 dB SPL; mode, 66.6), with slightly greater variance and range than used in the preschool setting. The environment with adults environment had the lowest mean and variation.

Figure 4 illustrates the results of dB SPL usage between the four communication environments using the primary statistics from Table 2 (mean, median, mode, and inter-quartile range). Generally, there is an increase in dB SPL parameters from environment to environment with the *free-play* time being the exception. The time in the *adult* environment was the lowest in terms of mean and median.

Voice Use Profile

The child's recorded F_0 and dB SPL were plotted in Figure 5 in terms of a Voice Use Profile, or a contour plot showing the phonation density in terms of F_0 and dB SPL. The Voice Use Profile (VUP) is similar to the elevation plot of a mountain: the closer the lines are together, the steeper the slope and, thus, the less dB range needed to result in a large frequency change. In Figure 5, the peak density was circled with a thick ellipse fit to capture the most frequently occurring 10% of all phonations. The area of the ellipse was calculated (in terms of dB Hz). To show the relation between F_0 and dB SPL for this most frequently occurring voice use, the slope of the major axis of the ellipse (in terms of dB/Hz) was calculated. Table 3 lists some general recording statistics as well as the area of these ellipses and the slope of the axis. The slope of the major axis of the ellipse indicates the general relationship between voice F_0 and vocal dB SPL for the most frequently occurring 10% of all vocalizations. While the VUP area and slope indicate how the child produced voice, it

does not illustrate the amount of voicing. Therefore, in order to associate the amount of voicing when comparing the different environments, voice percentage (percent of observation period that voicing occurred) was also listed in the table.

The *free-play* communication environment had the largest VUP area for the most frequently occurring 10% of all phonations (369 dB Hz). This area was nearly three times the area of the *adult* communication environment (121 dB Hz), where phonations were much more controlled. The other two environments were nearer to the *free-play* environment (*home* = 331 dB Hz, *preschool* = 322 dB Hz). Such areas indicate the most frequently used range of voice for the child. Voicing percentage was highest in the *home* and *free-play* environments (~23%) with the *adult* environment the lowest (~16%).

Figure 6 graphically presents the percent voicing, the VUP area (dB Hz), and the slope (Hz / dB SPL) from Table 3. In order to graphically depict these on the same figure, the VUP areas are shown at a tenth the actual value, with the actual value labeled next to the corresponding bar. The three metrics in Figure 6 show some relationship between the settings (going left to right), with the *home* metrics all higher than *adult* environment; the *preschool* all lower than *home*, and *free-play* all higher than *preschool*. While the relationship between these environments is not strong enough to show an increase in all the metrics going from the lowest to the highest, it does indicate perhaps some principle component underlying the child's voice use, or an underlying relationship between voicing percentage, voicing area, and Hz / dB slope.

DISCUSSION

To better understand the range of voice production of a child, and how a child may use her/his voice in different speaking environments, the above results are discussed for each of the four environments. In doing so, distinct differences and trends become evident.

Adult

A pattern emerged from the F_0 values collected in the *adult* environment (Figure 3.a). During this period, the child accompanied the investigator to an office/laboratory and was around only adults for 8 hours. Although the environment was very familiar to the child and he was around adults who were familiar to him, his vocal behavior suggests that he was aware of the difference in the environment and adjusted to it. Specifically, his F_0 for that day was lower than what was seen during a typical day for him (e.g., *preschool*, *home* below) with a more restricted range of values.

Of more particular importance, his F_0 during *adult* was more characteristic of values reported in the literature for his age and was also more similar to the results of the controlled tasks in Hunter (2009). Thus, it is likely that an environment with only adults and limited opportunities for play appeared to (at least in part) cause the child to alter his voice production, using lower F_0 values and more restricted inter-quartile range. It is also likely that this difference is, in part, related to fewer non-speech vocalizations (e.g., shrieks, squeaks, grunts) which would not be as likely to occur in an adult environment. It is also possible that in a controlled adult communication environment, a child may try to imitate the lower F_0 values he hears, or perhaps copy the speaking patterns of the adults, in an effort to communicate more effectively. While the current literature does not give a clear indication of whether this is a predictable effect, some early theories suggest that infants adjust their pitch towards caregivers (Lewis, 1936). Although two controlled studies found no such evidence (Siegel, Cooper, Morgan & Brenneise-Sarshad, 1990; McRoberts & Best, 1997), neither of these studies used older children nor collected long-term unsolicited data.

As would be expected, the distribution was more peaked when the child was with the adults than when he was in situations where opportunities for play were more frequent (kurtosis of 4.2 versus 3.0, Table 1). This relative lack of frequency range in the constrained communication environment suggests that a child in an adult setting may not use the same range as might be used in other settings (see below).

The child used the lowest dB SPL in this setting, confirming that the child spoke more softly when he was in an environment with only adults. The distribution of dB SPL for this environment appeared the most normal in shape. This environment also resulted in the smallest observed voicing percentage (16%), indicating an environment where the child was either sitting quietly, listening and observing, or otherwise preoccupied with quiet tasks. With the absence of play time with other children where more spontaneous speaking would occur, this is not unexpected. The *adult* environment also resulted in the smallest VUP area, again illustrating the reduced voicing range (both in dB SPL and F_0).

Home

Fundamental frequency values for the *home* communication environment were somewhat comparable to *preschool* (below), as seen in Figure 4. This was not unexpected given their general similarities: [1] both included some free playtime with peers or younger siblings; [2] both included some interactions with adults; [3] both were partly regulated by rules related to the type of voice used (i.e., “inside voice” in a classroom or home); and [4] both had some controlled, quieter time. Nevertheless, the child had a higher mode F_0 at *home*, perhaps indicating proportionally less time when the child was expected to use a more restrained voice. The slightly lower median and mean F_0 may have been indicative of fewer instances of excited free play, which result in fewer opportunities for the shrieks and squeals he would have uttered when he was with his peers at *preschool*.

The child's use of dB SPL in the *home* environment (looking at the distribution plots) was midway between the *adult* and *preschool* distributions, with the *adult* peaking lower and the *preschool* peaking higher. This may be because voice use during *home* time would not only include louder discussions and vocal competition with siblings, but also quieter dialogues with another family member.

The voicing percentage at *home* was the highest of the communication settings, similar to, and just above, the *free-play*. The values of area (dB Hz) and slope (rate of change of Hz per dB) in the VUP were between the values of the other settings.

Preschool

Like the *home* environment, the child used a wide range of frequencies in the *preschool* environment, with a peak around 280 Hz. Nevertheless, the instances of higher frequencies after the mode did not decrease as quickly as the *home* environment and there were other apparent differences. As mentioned, the *preschool* communication environment had a lower mode F_0 , but a higher median and mean, than the time at *home*. This deviation of median and mean from the mode indicate that the child spent more vocal time with high frequencies. In a classroom environment, the child would have a great deal of regulated instruction, so would not be excited enough to use a frequent higher F_0 that might be the result of squeals or screams (thus, F_0 mode would be lower). This tendency may be similar to the communication influences of the *adult* environment. On the other hand, the child was surrounded by peers, rather than just adults, who also have higher F_0 values. Further, during preschool or younger elementary school years, there is still enough play time with friends (and, thus, more excited utterances similar to the *free-play* environment discussed below) to skew the mean and median F_0 values higher (as can be seen in the distribution plots, Fig. 1).

During *preschool*, the child also used wide range of vocal levels (longest ellipse of the environments) and the highest dB SPL average, median and mode. This kind of voice use is not unexpected, even within a controlled classroom setting, where it is appropriate to answer questions and engage in discussions in a loud, projected voice, and less common to have quiet conversations. Further, a louder voice may also stem from the child's desire to project knowledge and confidence in front of his teacher and classmates. While the differences in average, median and mode were not more than a dB or two compared to two of the other environments, given the hours of recording contributing to the calculation, the differences are noteworthy.

Looking at the VUP in Figure 5, the *preschool* communication environment had the least frequency change in semitone per dB (0.55 ST/dB, Table 3). Its shape was the most unique, being a long narrow ellipse. Looking at the larger picture of the VUP and not the top 10%, the *preschool* environment seems to have a curved trajectory or a combination of two different slopes, with a steep slope for low F_0 and dB SPL values, and a very shallow slope at high F_0 and dB SPL. Thus, the child used a wide range of loudness levels when he used low frequencies. On the other hand, when the child used a wide range of higher frequencies, he used a relatively small range of high vocal levels. These results may suggest that when the child speaks with a lower pitch (possibly during class) he speaks louder without a great deal of pitch variation. In the higher frequencies (possibly at play), he uses a wide range of frequencies at a high level.

Free-Play

The child used a generally higher and wider range of F_0 values for all the metrics while playing with his friends in the *free-play* environment than he did in the other environments (Figure 4). Also, unlike the other environments, no distinct peak (mode) was seen in Figure 1, suggesting nearly equal instances of voice frequencies over the higher vocal range. This is indicated by the largest inter-quartile range, which likely reflected the large number of higher vocalizations such as the squeals and screams that could occur along with normal speech when playing on the gym equipment. Finally, the child had the highest occurrences of the lowest F_0 values during this time, perhaps from low grunts and growls which may also occur during this type of play.

The *free-play* environment had the widest variation in loudness. Proportionally, he used the loudest voice production of all the environments as indicated by the interquartile range, which was expected given the loud voice he likely used during play. However, the child's mean and median dB SPL was also both less intense on average and most frequently less intense than the *home* and *preschool* environments (just 2 dB louder than in the *adult* environment). One possible explanation for this is that in the *free-play* environment, the child engaged in softer, isolated non-verbal vocalizations (i.e., grunts when climbing and running) more often than usual. It is also possible that the child periodically tried to whisper during isolated conversations with a playmate, and such whispering attempts in children often contain low-level phonation. Such conversations would be unacceptable in a classroom environment and less expected in the other two communication environments.

The *free-play* communication environment was unique with the most F_0 change per dB (nearly 1 semitone for every dB). Even with the likely production of squeaks and squeals, the child was not abnormally vocally active when compared to the other environments, having nearly the same voice activity percentage as the *home* environment, though the phonation types were quite different.

SUMMARY AND CONCLUSIONS

The current study analyzed a single child's natural voice data collected over several days, to provide real-world information about voice use outside of a controlled research situation. The child's fundamental frequency (F_0) and voice level (dB SPL) were captured in four different communication environments. Relationships between F_0 and dB SPL for the environments were also quantified using a Voice Use Profile (VUP). The implication of such data could be beneficial to understanding voice use (and abuse) leading to the frequent occurrences of such child pathologies as vocal nodules (Akif et al., 2004), as well as social communication behavior as children learn to understand and adapt to their environment.

The child in this study produced significantly different F_0 values dependent on the type of communication environment. The unrestrained *free-play* environment produced the largest ranges of both F_0 and dB SPL. In contrast, the time the child spent with adults in the office, the *adult* environment, resulted in the most limited range of F_0 and dB SPL.

Unique to this study, the child's voice data in the different environments were presented as Voice Use Profiles (VUP). The VUP shows the actual use of the voice in terms of F_0 and dB SPL, the relationship between F_0 and dB SPL, and an analysis of the most frequently used parts of the voice range. The results indicate that the child's vocal level, and vocal pitch in relation to vocal loudness, is most similar to an adult's pattern when the child is in the *preschool* speaking environment.

As suggested in Hunter (2009), common statistics should not be automatically used without understanding the distribution of the values. The statistical mean may not accurately capture child voice usage in natural communication environments because the vocal level distribution is often not normal. Mode is useful for understanding the most frequently occurring values, particularly in the case of identifying the most characteristic voice F_0 . For example, it was demonstrated that the child most frequently phonated at 290 Hz in the *Home* environment but the average frequency in that same environment was 379 Hz. Mode, however, may not be informative in environments where there is a broad range of F_0 and dB SPL values, or where there is no clear peak in the distribution, such as the *free-play* environment in the current study. Finally, while standard deviation values may be adequate when there is a normal distribution (e.g., dB SPL were nearly normal), interquartile range is a more robust indicator of range because it can quantify both a skewed and a normal distribution.

The results of this study, though a single-subject case study, provide valuable preliminary data. However, there are limitations to this study which suggest avenues for future research. This study does not provide a complete picture of how all children use their voices in various environments because every child will most likely be different and it is not yet known how much variation would be seen. In addition, the communication environments presented here would be difficult to repeat. Thus, it would be valuable to conduct future studies with multiple participants to better capture the statistical and clinical significance of these findings, as well as the size of the effect. Further, by expanding the population pool, other potentially important speaking environments can be examined, including interactions between children and voice/speech clinicians.

Another possible limitation of the current study is the unequal time for data collected across the different environments, which may have impacted the comparison. Future studies with more data could relate specific voicing characteristics to the variations in the environments. For example, the current study treated the entire *preschool* environment as one type of environment, although distinctions between class time and recess time would be of interest.

Finally, it is likely that the comparison of environments (e.g., *free-play* vs. *adult*) with differing ratios of speech versus non-speech vocalizations (e.g. screams, squeals and laughter) may skew the results. The instrumentation in the current study only captures phonation during speech, missing voiceless speech production, and is not able to differentiate between speech and non-speech vocalizations. Future studies could use other devices or transducers (such as throat microphones, or binaural microphones with adaptive canceling) to monitor the full speech range. Nevertheless, it is important to note that the dosimeter used in the current study is simpler than these other options and maintains the privacy of the subjects because speech cannot be recreated from the recording.

Because the four communication environments described here each contain several variables which may have affected the outcomes, future studies should isolate aspects of individual environments such as location (e.g., size of room, noise level in the environment, or inside/outside condition), familiarity of the child with the communication partner(s) (e.g., family member, relative, friend, or new associate), level of perceived authority (e.g., parent, teacher, adult, older child, or younger child), or the number and age(s) of the communication partner(s). If characteristics of a child's environment were isolated, the impact of each of these conditions individually on F_0 and dB SPL could be examined in more detail. In addition, further studies should investigate speech compared to non-speech vocalizations (yell, screams, and squeals), including the ratio of speech vs. non-speech behaviors and the potential impact of that ratio on voice disorders, such as the risk of pediatric vocal abuse.

Based on the current study, it appears that communication environment can significantly affect a child's patterns of voice use. This study also highlights the limitations of evaluating a child and developing a comprehensive treatment plan when the information is based solely on behaviors observed in the speech treatment or clinic room. Researchers should be aware that while the observation/recording environment may elicit results that may be repeatable, the results may still not represent a child's natural voice use. Thus, researchers and speech therapists working in the schools should, therefore, include parental reports of children in a variety of natural communication environments in order to obtain the most accurate representation of their typical voice and speech patterns. If possible, observations during class or play environment would further clarify the child's natural voice use. Finally, the child's tolerance of the voice dosimeter (as well as its durability) in collecting multiple hours of data indicate the potential use of similar devices for obtaining "real world" voice data in the future.

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 authors would like to thank the research team (both past and present) at the NCVS with many supporting roles (Dosimeter Team: Ingo Titze, Jan Svec, Peter Popolo, Andrew Starr, Albert Worley); and to Laura M. Hunter for literature search and technical review of the document.

REFERENCES

- Akif KM, Okur E, Yildirim I, Guzelsoy G. The prevalence of vocal fold nodules in school age children. *International Journal of Pediatric Otorhinolaryngology*. 2004; 68(4):409–412. [PubMed: 15013605]
- Awan SN, Mueller PB. Speaking fundamental frequency characteristics of white, African American, and Hispanic kindergartners. *Journal of Speech and Hearing Research*. 1996; 39:573–577. [PubMed: 8783135]

- Baker S, Weinrich B, Bevington M, Schroth K, Schroeder E. The effect of task type on fundamental frequency in children. *International Journal of Pediatric Otorhinolaryngology*. 2008; 72:885–889. [PubMed: 18395803]
- Gramming P, Sundberg J, Ternstrom S, Leanderson R, Perkins W. Relationship between changes in voice pitch and loudness. *Journal of Voice*. 1988; 2(2):118–126.
- Hacki T, Heitmuller S. Development of the child's voice: premutation, mutation. *International Journal of Pediatric Otorhinolaryngology*. 1999; 49(Suppl 1):S141–S144. [PubMed: 10577793]
- 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]
- Heylen L, Wuyts FL, Mertens F, De Bodt M, Pattyn J, Croux C, Van de Heyning PH. Evaluation of the vocal performance of children using a voice range profile index. *Journal of Speech, Language, & Hearing Research*. 1998; 41(2):232–238.
- Hunter EJ. A comparison of a child's fundamental frequencies in structured elicited vocalizations versus unstructured natural vocalizations: A case study. *International Journal of Pediatric Otorhinolaryngology*. 2009 Apr; 73(4):561–571. [PubMed: 19185926]
- Lewis, MM. *Infant Speech: A Study of the Beginnings of Language*, Harcourt. New York: Brace and Co; 1936.
- McRoberts GW, Best CT. Accommodation in mean f0 during mother-infant and father-infant vocal interactions: a longitudinal case study. *Journal of Child Language*. 1997; 24:719–736. [PubMed: 9519592]
- 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.
- Siegel GM, Cooper M, Morgan JL, Brenneise-Sarshad R. Imitation of intonation by infants. *Journal of Speech, Language, and Hearing Research*. 1990; 33:9–15.
- Svec, JG.; Hunter, EJ.; Popolo, PS.; Rogge-Miller, K.; Titze, IR. NCVS Memo No 02. The Calibration and Setup of the NCVS Dosimeter. 2004. (Rep. No. April 2004), <http://www.ncvs.org/e-learning/technical.html> (date last viewed 01/06/2011).
- Svec JG, Popolo PS, Titze IR. Measurement of vocal doses in speech: experimental procedure and signal processing. *Logopedics Phoniatrics Vocology*. 2003; 28:181–192.
- Svec JG, Titze IR, Popolo PS. Estimation of sound pressure levels of voiced speech from skin vibration of the neck. *Journal of the Acoustical Society of America*. 2005; 117(3, Pt.1):1386–1394. [PubMed: 15807026]
- Titze IR, Svec JG, Popolo PS. Vocal dose measures: quantifying accumulated vibration exposure in vocal fold tissues. *Journal of Speech, Language, and Hearing Research*. 2003; 46:919–932.
- Wuyts FL, Heylen L, Mertens F, Du CM, Rooman R, Van de Heyning PH, De Bodt M. Effects of age, sex, and disorder on voice range profile characteristics of 230 children. *Annals of Otology, Rhinology, Laryngology*. 2003; 112(6):540–548.

Figure 1.

The F_0 distributions from four different environments: [a] with *adults*; [b] at *preschool* [c] at *home*; and [d] at *free-play* with other children.

Figure 2. Mean, median, Quartile 1, Quartile 3, and mode for F₀ in the four communication environments (*Adult, Home, Preschool, Free-play*). The Quartiles are represented as error bars around the median

Figure 3.

The Level (estimated dB SPL @ 30 cm) distributions from four different environments: [a] with *adult*; [b] at *preschool*; [c] at *home*; and [d] at *free-play* with other children.

Figure 4. Mean, median, Quartile 1, Quartile 3, and mode for the four communication environments (*Adult, Home, Preschool, Free-play* with other children). The quartiles are represented as error bars around the median.

Figure 5.

The Voice Use Profile from four different environments: [a] with *adults*; [b] *at home*; [c] at *preschool*; and [d] at *free-play* with other children. The VUP illustrates phonation density as contour plots. An ellipse was fitted to the most frequent 10% of phonations.

Figure 6.

Percent voicing, VUP area (Hz dB) and slope (Hz / dB) are shown for the four communication environments (*Adult, Home, Preschool, Free-play*). Note that the VUP area (the ellipse from Figure 5) values have been divided by 10 in the figure for overall scaling purposes, but the actual values are shown above the bars.

Table I

F_0 statistics for four different communication environments (with *Adults*, at *Home*, at *Preschool*, at *Free-play* with other children).

	Adult	Home	Preschool	Free-play
Mean:	334.4	378.7	396.3	423.4
Variance:	8057.8	13298.0	15150.0	19087.0
Standard Deviation:	89.8	115.3	123.1	138.2
1 st Quartile (25 th percentile):	269.2	290.7	301.5	323.0
2 nd Quartile (Median):	312.2	355.3	376.8	409.1
3 rd Quartile (75 th percentile):	376.8	452.2	463.0	495.3
Inter-Quartile Range	107.6	161.5	161.5	172.3
Mode:	258.4	290.7	279.9	366.1
Kurtosis:	4.2	3.9	3.3	3.0

Table II

Estimated dB SPL (30 cm) statistics for four different communication environments (with *Adults*, at *Home*, at *Preschool*, at *Free-play* with other children).

	Adult	Home	Preschool	Free-play
Mean:	64.1	66.2	67.0	65.6
Variance:	32.2	38.7	35.5	46.1
Standard Deviation:	5.8	6.2	5.9	6.8
1 st Quartile (25 th percentile):	60.0	61.6	63.0	60.0
2 nd Quartile (Median):	63.0	65.6	67.0	65.0
3 rd Quartile (75 th percentile):	67.0	70.6	71.0	70.0
Inter-Quartile Range	7.0	9.0	8.0	10.0
Mode:	63.0	66.6	68.0	63.0
Kurtosis:	2.6	2.4	2.3	2.3

Table III

Various recording statistics and Voice Use Profile (area and slope) of the most frequent 10% of all phonations on a map of dB and Hz. Final slope in semitone per dB [$ST=39.86 \log_{10}(F_0/16.25)$].

	Adult	Home	Preschool	Free-play
Voicing Percentage	15.9	23.4	20.6	23.3
Hours Recorded	5.6	10.7	3.0	2.5
N data points	107324	298204	74311	69771
Voicing area (dB $\times$ Hz)	121	331	322	369
Slope of Major Axis (dB/Hz)	0.088	0.074	0.120	0.055
Slope (Hz/dB)	11.3	13.6	8.3	18.2
Slope of Major Axis (ST/dB)	0.741	0.889	0.546	1.190