Voice change as a new measure of male pubertal timing: A study among Bolivian adolescents

Carolyn R. Hodges-Simeon, Michael Gurven, Rodrigo A. Cárdenas & Steven J. C. Gaulin

1Department of Anthropology, University of California, Santa Barbara, CA 93106, USA, and 2Department of Psychology, Pennsylvania State University, University Park, PA 16802, USA

Background: Age at menarche is often used to measure maturational tempo in girls. Unfortunately, no parallel marker exists for boys. It is suggested that voice change has a number of advantages as a marker of the timing and degree of male pubertal development.

Aim: Traditional auxological methods are applied to voice change in order to compare differential development both between (males vs females; Tsimane vs North American; better vs worse condition) and within (voice vs height; fundamental frequency vs formant structure) populations.

Subjects and methods: Fundamental and formant frequencies, as well as height and weight, were measured for 172 Tsimane males and females, aged 8–23. Participants were assigned to ‘better’ or ‘worse’ condition based on a median split of height-for-age and weight-for-age z-scores.

Results: Results support dramatic vocal changes in males. Peak voice change among Tsimane male adolescents occurs ~1 year later than in an age-matched North American sample. Achieved adult male voices are also higher in the Tsimane. Tsimane males in worse condition experience voice change more than 1 year later than Tsimane males in better condition.

Conclusion: Voice change has a number of attractive features as a marker of male pubertal timing including its methodological and technical simplicity as well as its social salience to group members.

Keywords: Puberty, adolescence, voice, fundamental frequency, formants, formant dispersion, peak height velocity, pitch

INTRODUCTION

From a life-history perspective, sexual maturation marks a strategic shift involving a re-allocation of resources away from growth and towards reproduction (Stearns 1992; Hill 1993; Kaplan et al. 2000). Phenotypic (e.g. endocrine, somatic) changes accompany dramatic alterations in social, economic and reproductive roles and relationships. Differential timing of the onset and duration of these changes is thought to occur in response to variation in social and ecological inputs (Wyshak and Frisch 1982; Susanne 1985; Evelthe and Tanner 1990; Belsky et al. 2007; Bolen 1999; Cole 2003; Parent et al. 2003; Walker et al. 2006; Cameron 2007). As the study of development progressively incorporates this theoretical perspective, hypothesis testing about both evolutionary and proximate (internal and external) environmental triggers will depend on multiple reliable measures of this important biosocial transition.

The primary measure of the timing and pace of reproductive maturation has been the age at menarche (Wyshak and Frisch 1982; Bolen 1999; Fredriks et al. 2000; Cole 2003; Herman-Giddens 2006; Walker et al. 2006; McDowell et al. 2007). This defining event has been occurring earlier for at least the past century (Wyshak and Frisch 1982) and continues to show declines of ~1 month per decade in the US and Western Europe (Fredriks et al. 2000; Herman-Giddens 2006; Ong et al. 2006; McDowell et al. 2007). Compared with more developed countries, less developed countries have a later age at menarche (Bolen 1999; Walker et al. 2006), which often falls as living conditions improve (Hoshi and Kouchi 1981; Bolen 1999; Parent et al. 2003). Studies of menarcheal age have provided important breakthroughs in understanding the impact of social and environmental factors on reproductive development (e.g. Belsky et al. 2007; Walker et al. 2006). Unfortunately, parallel investigations into male pubertal timing have been more difficult because no equivalent marker of masculine pubertal development has been identified.

Because the pubertal period differs so markedly for the sexes (Tanner 1981), markers of male reproductive development are needed. Compared to females, males reach puberty later (Tanner 1981; Bolen 1999), develop divergent
primary and secondary sexual characteristics (Tanner 1981) and employ different life history and reproductive strategies (Trivers 1972; Kaplan et al. 2000). In addition, growth and development among males may be more sensitive to environmental hardship (Stinson 1985; Cole 2000), such that supportive or adverse conditions impact their development more strongly than for females (Kuh et al. 1999). The timing of reproductive landmarks also differs: females develop secondary sexual characteristics prior to adult reproductive functioning, whereas, for males, the opposite pattern prevails (Bogin 1999).

Given the growing focus on sexual maturation as a major life-history transition, more precise markers of male reproductive development could support theory formulation and testing. Conventional measures include the age at peak height velocity (Eveleth and Tanner 1990; Bogin 1999), testicular volume (Harries et al. 1997; Campbell et al. 2004), Tanner stages of genital and pubic hair development (Harries et al. 1997; Bogin 1999) and endocrine biomarkers (e.g. DHEA-S and testosterone; Halpern et al. 1998; Ellis and Essex 2007). Unfortunately none of these measures is ideal. Although height itself has relatively high reliability, peak height velocity can sometimes be difficult to detect in cross-sectional or energetically stressed samples (Bogin 1999). Tanner stages are widely used but are somewhat subjective; as such, measures collected by different researchers may be incomparable (Brooks-Gunn and Warren 1985; Herman-Giddens 2006) and provide a poor basis for population comparisons. Both Tanner stages and testicular volume often require the presence of a physician or trained health professional, making them sub-optimal measures for sensitive or reticent populations or for studies with limited funding (Coleman and Coleman 2002). Endocrine biomarkers provide a desirable level of objectivity and reliability (McCade et al. 2007), but the expense and difficulty associated with storage, transport and analysis discourages their widespread adoption, especially by field researchers.

Voice frequency, which changes dramatically during adolescence (Hollien et al. 1994; Harries et al. 1997, 1998; Huber et al. 1999), is an under-utilized male pubertal marker. With digital means of collection, storage and analysis (e.g. Boersma and Weenik 2010), acoustic samples can be easily gathered in a wide variety of settings, without extensive technical training. For these pragmatic reasons, male vocal characteristics could be useful in studies of within- and between-population comparisons of peri-pubertal timing and a candidate marker for the examination of secular trends or epidemiological factors in male puberty (Pedersen et al. 1986; Juul et al. 2007). Voice frequency would be an especially useful marker of male pubertal development if male vocal changes were relative precipitous compared to other developmental transitions.

Like height, the voice is a socially and economically significant characteristic, particularly for males. Research indicates that a lower voice impacts social perceptions, including assessments of age (Collins 2000; Feinberg et al. 2005), size (Collins 2000; Feinberg et al. 2005), physical and social dominance (Puts et al. 2006, 2007; Wolff and Puts 2010) and attractiveness to women (Collins 2000; Feinberg et al. 2005; Puts 2005). Men with lower voices had greater mating success in a US sample (Puts 2005; Hodges-Simeon et al. 2011) and greater reproductive success in a sample of Hadza hunter-gatherers (Apicella et al. 2007). Thus, unlike pubertal indicators such as testicular volume, manifest changes in males’ voices are likely to coincide with social and economic changes in adolescents’ lives—a transition that would be of interest to those studying biocultural dimensions of human puberty and adolescence.

Although male voice change from ‘high’ to ‘low’ is self-evident to listeners, acoustic dimensions must be measured objectively and precisely in order to be useful to researchers. The perceptual experience of a ‘low’ voice is the psycho-physical product of two acoustic components: fundamental frequency (a perceptual correlate of ‘pitch’; $F_0$) and formant structure (Puts et al. 2007; Feinberg et al. 2011). $F_0$ is a product of the length, tension and cross-sectional area of the vocal folds, such that larger vocal folds produce lower $F_0$ (Williams and Eccles 1990; Harries et al. 1998). Speakers can vary $F_0$ to some extent by changing the tension of the vocal folds and they often do so for prosodic purposes. However, stable individual differences exist in the baseline $F_0$ as a result of the length and cross-sectional area of the folds (Titze 1994).

Formant dispersion ($D_F$, Fitch 1997) and formant position ($P_F$, Puts et al. 2012) are measures of the spacing of the formants—energy peaks in the harmonic spectrum (Fitch 1997). Lower and more closely spaced formants give a more ‘resonant’ quality to the voice—analogous to the fuller and deeper resonance of a middle C played on a cello vs a violin. Formant structure is a function of the length and shape of the vocal tract (or the depth of the larynx in the throat), which jointly define the region that sound waves traverse from the vocal folds through the oral cavity (Baken 1987; Titze 1994). Thus, differences in acoustic dimensions reflect differences in the biology of the vocal tract. See Figure 1 for a visual representation of formant frequencies in a pre- and post-pubertal male. A longer vocal tract in the older male generates lower and more closely spaced formants; thus, post-pubertal formants are lower. In the present study, we report two measures of formant structure: $D_F$ and $P_F$ (Puts et al. 2012).

In both sexes, $F_0$ drops steadily throughout childhood as the growth of the vocal folds scales with overall body growth (Vuorenkoski et al. 1978). However, working through androgen receptors on the larynx (Newman et al. 2000), testosterone lengthens and thickens the vocal folds in males (Kahane 1978; Jenkins 1998), causing a rapid drop in $F_0$ during Tanner stages 3 and 4, usually around 12–15 years old (Hollien et al. 1994; Harries et al. 1997, 1998; Huber et al. 1999). Male $F_0$ continues to lower more gradually throughout the 20s and 30s, but rises again in old age (Baken 1987). Similarly, vocal tract length scales with body size growth, with little sexual dimorphism prior to puberty (Vorperian et al. 2005). At puberty, however, males’ larynges begin to descend further than females (Titze 1994).
The general goals of the present study are to highlight voice change as an important aspect of adolescent male biology and to suggest ways that change in $F_0$ and formant structure can be useful as an additional measure of the peri-pubertal life-history shift from growth to reproduction. In addition to age-specific means and standard deviations, we use traditional auxological tools to capture the ‘spurt’ in adolescent voice change: distance, velocity and acceleration curves and their attendant features, age at peak velocity, minimal velocity (or age at take-off velocity), maximal acceleration and deceleration (Bogin 1999; Cameron 2004). We then use these measures to make several comparisons of voice change among adolescent males and females in an energetically stressed population, the Tsimane of Bolivia. Several studies have previously described aspects of adolescent voice change (Hollien et al. 1994; Harries et al. 1997, 1998; Huber et al. 1999); however, this is the first study to examine voice change: (a) using auxological tools and (b) in a non-westernized population living under considerable immune and energetic stress, where growth and development are often adversely affected (Susanne 1985; Eveleth and Tanner 1990; Bogin 1999; Cole 2003).

Within the framework of these general goals we have five specific objectives:

- First, both male and female $F_0$ and formant structure are modelled in order to highlight the pronounced difference between male and female voice change and to identify the ages at which the sexes diverge on these vocal parameters.
- Second, we compare the age of peak $F_0$ change between Tsimane and US adolescent boys, where the latter estimate derives from published North American data (Huber et al. 1999). Other indicators of pubertal timing (i.e. age at menarche, age at peak height velocity) suggest delayed development in our Bolivian population (Walker et al. 2006); therefore, we predict that Tsimane development will be delayed on the vocal dimension as well.
- Third, male voice change is modelled separately for boys in better vs worse condition, as defined by weight-for-age and height-for-age $z$-scores (Kuczmarski et al. 2000). Within the Tsimane, variation is present in access to food, healthcare, clean water and other variables that affect development (Gurven et al. 2007). We hypothesize that peak voice change will occur later for males in worse condition than for males in better condition.
- Fourth, post-pubertal values of $F_0$ and $D_f$ are compared to published averages from the US and Europe. Just as a slower rate of linear growth is associated with shorter stature among those in relatively worse condition (Walker et al. 2006; Godoy et al. 2010), we predict that a later age at peak $F_0$ and $D_f$ velocity combined with a slower rate of change at all ages will result in higher voices among post-pubertal Tsimane males.
- Finally, we compare our measures of $F_0$ and formant structure change with another widely used pubertal marker—the linear growth spurt in height. Here we address two related questions. First, are vocal and stature changes temporally correlated? Second, is the pace of change in each similarly dispersed across adolescence or is one more temporally compressed than the other? This is important because the more ‘instantaneous’ the change (e.g. menarche), the more precise a measure of reproductive timing it can provide.

**MATERIALS AND METHODS**

**Population**

The Tsimane are South American forager-horticulturalists living in the Amazonian lowland forests of the Beni region of Bolivia. Villages are settled along the banks of the Maniqui River, between market towns San Borja and San Ignacio de Mojos. Several different villages were sampled for the present study. Tacuaral del Mato (population was ~ 553 in 2010, $n = 74$) and Campo Bello (pop. 328, $n = 56$)
are ~ 1 hour from San Borja. Jamanchi Uno (pop. 221, n = 42) is a smaller, less-acculturated village 2 hours from San Borja. The Tsimane participate in a broad, NIH/NIA-funded project (Tsimane’ Health and Life History Project [THLHP]; Co-PIs: Michael Gurven and Hillard Kaplan).

In this population, the average male adolescent growth spurt in stature lasts from age 13–18 years; the age at peak height velocity is 14 (age 12 for females). The overall rate of child–juvenile growth velocity is 5.2 cm/year (4.9 for females) until an adult size of 163 cm (149 cm for females) is reached. The average age at menarche is 13.9 according to one estimate from a mixed cohort sample (Walker et al. 2006); however, girls in the most recent cohort—especially those in more acculturated villages—may experience menarche earlier (McAllister et al. 2012). Self-reported age at menarche for the current sample of post-menarche girls was 13.0. Tsimane children and young adults experience high rates of endemic infections—primarily respiratory and gastrointestinal—and intestinal parasitism (Gurven et al. 2008). These infections likely account for the low height-for-age profiles of children and young adults (McDade et al. 2005). Life expectancy at age 15 is 43 additional years (Gurven et al. 2007).

Participants
Eighty-one females, aged 8–20, and 91 males, aged 8–23, participated in this IRB-approved study. The distribution across ages can be found in Table I.

Ages were obtained using two sources of information. First, each participant was asked his or her age and date of birth. Many participants (mostly under the age of 12; n = 69) did not know their exact birth date. One participant did not know his age or birth date. Second, stated age was compared with the THLHP census (see Gurven et al. 2007 for detailed age methods). This census information was primarily obtained from demographic interviews conducted over the past 10 years. Census ages are most accurate for younger ages (vs middle and older ages) because interviews were conducted with parents when their child was young. In addition, parents of the current participants are more likely to keep records of their child’s birth than earlier generations. Therefore, the census age was used when it was in conflict with stated age. When participants neither knew their birth date nor could be found in the census (n = 8), they were assigned their stated whole number age plus 0.5. Omitting these individuals did not alter any of the results reported here and so are included in the sample.

Acoustic recording
Voices were recorded using a Sony PCM-M10 digital audio recorder and an Audio-Technica lavaliere microphone. The microphone was mounted on an adjustable headset in order to standardize the distance from the lips and reduce background noise. Due to differing head sizes, the headset was adjusted for each person so that the microphone was ~ 5 centimetres from the lips. Voice samples were recorded in mono using a sampling rate of 44.1 kHz and 16-bit quantization. Each recording was saved as a high-quality uncompressed linear PCM.wav file.

Participants were shown five photographs representing easily recognized objects in Tsimane life (cat, dog, plantains, soccer ball and frog) and asked to name each object. These particular words were chosen because they terminate in five different vowel sounds (mishi: ‘ee’, achuji: ‘oo’, pe’re: ‘a’, perota: ‘ah’ and oco: ‘oh’).

Acoustic analysis
For all participants’ recordings, mean $F_0$ and formant structure ($D_1$ and $P_f$), all measured in Hertz (Hz), were determined using Praat voice analysis software (Version 5.13.7; Boersma and Weenik 2010).

Formants were obtained using methods described by Puts et al. (2012); however, formant ceilings were adjusted for different developmental groups. Participants completed the Pubertal Development Scale (PDS; Petersen et al. 1988), which has been shown to be a reliable approximation of the Tanner stages (Brooks-Gunn & Warren 1985; Petersen et al. 1988; Coleman and Coleman 2002). Males and females in stages 1 and 2 (mean age ≈ 10.9 ± 1.6) were measured using a formant ceiling of 7000 Hz. Females in stages 3, 4 and 5 (mean age ≈ 15.4 ± 2.7) and males in stage 3 (mean age ≈ 15.1 ± 1.5) were measured using a formant ceiling of 6000 Hz. Males in stages 4 and 5 (mean age ≈ 18.2 ± 2.3) were measured with a formant ceiling of 5000 Hz.

$F_1$ through $F_4$ were measured at each glottal pulse, targeting voiced speech only and then averaged for analyses (mean number of glottal pulses per file = 527 ± 273). In order to calculate $P_f$, the first four formants were standardized using between-sex means and standard deviations (Puts et al. 2012). The median was used for analyses. $D_f$ was calculated by taking the average of the distances between the first four formants (Fitch 1997). Because within-sample means are used to calculate standardized formants for $P_f$ it does not provide an optimal measure for between-study comparisons of formant frequencies. Therefore, we retain the more conventional $D_f$ as our primary measure of formant structure.

Height, weight and nutritional condition
Standard anthropometric procedures were used to collect height and weight (Lohman et al. 1988); participants were measured without shoes and in light clothing. Sex-specific weight-for-age (WAZ) and height-for-age (HAZ) z-scores were computed using CDC (2000) reference curves in EpilInf (version 3.5.3). WAZ and HAZ were averaged to create a composite condition variable. Participants were assigned to either ‘better’ ($M = -0.5$) or ‘worse’ ($M = -1.8$) condition based on a median split of this condition variable.

RESULTS
Exploration of male and female $F_0$ and formant structure ($P_f$ and $D_f$) at different ages
See Table I for mean values of $F_0$ (Hz; $M = 219.3$, SD = 50.2, Range: 106.8–353.4), $D_f$ (Hz; $M = 1356$, SD = 142, Range: 947–1664) and $P_f$ ($M = 1.17$, SD = 1.06, Range: −1.12

Annals of Human Biology
Table I. Mean and standard deviation of fundamental frequency ($F_0$), individual formants ($F_1$–$F_4$) and formant structure (formant dispersion [$D_f$] and formant position [$P_f$]) across adolescent ages.

<table>
<thead>
<tr>
<th></th>
<th>$\leq 10$</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>$\geq 19$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 48)</td>
<td>(n = 17)</td>
<td>(n = 18)</td>
<td>(n = 15)</td>
<td>(n = 20)</td>
<td>(n = 9)</td>
<td>(n = 12)</td>
<td>(n = 11)</td>
<td>(n = 10)</td>
<td></td>
</tr>
<tr>
<td>$F_0$ Males</td>
<td>243 ± 29</td>
<td>227 ± 27</td>
<td>242 ± 29</td>
<td>182 ± 30</td>
<td>166 ± 44</td>
<td>133 ± 14</td>
<td>123 ± 5</td>
<td>140 ± 5</td>
<td>125 ± 26</td>
<td>141 ± 33</td>
</tr>
<tr>
<td>Females</td>
<td>259 ± 38</td>
<td>244 ± 28</td>
<td>252 ± 40</td>
<td>227 ± 8</td>
<td>235 ± 18</td>
<td>219 ± 31</td>
<td>258 ± 46</td>
<td>221 ± 14</td>
<td>231 ± 27</td>
<td>242 ± 25</td>
</tr>
<tr>
<td>U</td>
<td>176</td>
<td>20</td>
<td>34.0</td>
<td>1**</td>
<td>9**</td>
<td>0**</td>
<td>0</td>
<td>0*</td>
<td>0*</td>
<td>0**</td>
</tr>
<tr>
<td>$D_f$ Males</td>
<td>1474 ± 78</td>
<td>1492 ± 60</td>
<td>1387 ± 96</td>
<td>1354 ± 31</td>
<td>1304 ± 218</td>
<td>1251 ± 84</td>
<td>1199 ± 124</td>
<td>1226 ± 113</td>
<td>1119 ± 98</td>
<td>1083 ± 80</td>
</tr>
<tr>
<td>Females</td>
<td>1470 ± 97</td>
<td>1414 ± 96</td>
<td>1379 ± 109</td>
<td>1309 ± 94</td>
<td>1270 ± 51</td>
<td>1270 ± 69</td>
<td>1228 ± 67</td>
<td>1274 ± 66</td>
<td>1302 ± 54</td>
<td>1337 ± 41</td>
</tr>
<tr>
<td>U</td>
<td>237</td>
<td>22</td>
<td>33</td>
<td>16</td>
<td>44</td>
<td>13</td>
<td>6</td>
<td>0*</td>
<td>0*</td>
<td>0**</td>
</tr>
<tr>
<td>$P_f$ Males</td>
<td>2.09 ± 0.43</td>
<td>2.06 ± 0.21</td>
<td>1.51 ± 0.76</td>
<td>0.99 ± 0.51</td>
<td>0.55 ± 1.05</td>
<td>-0.10 ± 0.53</td>
<td>-0.22 ± 0.88</td>
<td>-0.29 ± 0.56</td>
<td>-0.62 ± 0.71</td>
<td>-0.48 ± 0.65</td>
</tr>
<tr>
<td>Females</td>
<td>2.33 ± 0.75</td>
<td>1.79 ± 0.56</td>
<td>0.94 ± 0.83</td>
<td>0.99 ± 0.55</td>
<td>0.45 ± 0.31</td>
<td>0.50 ± 0.46</td>
<td>0.16 ± 0.88</td>
<td>0.54 ± 0.50</td>
<td>0.91 ± 0.32</td>
<td>0.44 ± 0.43</td>
</tr>
<tr>
<td>U</td>
<td>190</td>
<td>25</td>
<td>20</td>
<td>22</td>
<td>44</td>
<td>6</td>
<td>4</td>
<td>2*</td>
<td>0*</td>
<td>5*</td>
</tr>
<tr>
<td>$F_1$ Males</td>
<td>472 ± 43</td>
<td>493 ± 48</td>
<td>463 ± 68</td>
<td>428 ± 82</td>
<td>433 ± 27</td>
<td>362 ± 24</td>
<td>427 ± 54</td>
<td>403 ± 19</td>
<td>390 ± 29</td>
<td>439 ± 45</td>
</tr>
<tr>
<td>Females</td>
<td>486 ± 60</td>
<td>488 ± 50</td>
<td>412 ± 74</td>
<td>445 ± 38</td>
<td>424 ± 48</td>
<td>429 ± 25</td>
<td>415 ± 50</td>
<td>417 ± 59</td>
<td>459 ± 55</td>
<td>380 ± 85</td>
</tr>
<tr>
<td>U</td>
<td>185</td>
<td>31</td>
<td>23</td>
<td>20</td>
<td>46</td>
<td>1**</td>
<td>3</td>
<td>7</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>$F_2$ Males</td>
<td>2261 ± 149</td>
<td>2148 ± 163</td>
<td>2124 ± 228</td>
<td>2021 ± 216</td>
<td>1874 ± 260</td>
<td>1897 ± 134</td>
<td>1883 ± 277</td>
<td>1684 ± 249</td>
<td>1715 ± 114</td>
<td>1655 ± 75</td>
</tr>
<tr>
<td>U</td>
<td>171</td>
<td>30</td>
<td>23</td>
<td>19</td>
<td>46</td>
<td>14</td>
<td>2</td>
<td>5</td>
<td>0**</td>
<td>0**</td>
</tr>
<tr>
<td>$F_3$ Males</td>
<td>3677 ± 170</td>
<td>3613 ± 170</td>
<td>3471 ± 139</td>
<td>3322 ± 127</td>
<td>3106 ± 341</td>
<td>2921 ± 187</td>
<td>3007 ± 216</td>
<td>2788 ± 195</td>
<td>2713 ± 230</td>
<td>2732 ± 182</td>
</tr>
<tr>
<td>Females</td>
<td>3746 ± 232</td>
<td>3738 ± 207</td>
<td>3317 ± 201</td>
<td>3370 ± 281</td>
<td>3159 ± 192</td>
<td>3142 ± 168</td>
<td>3040 ± 88</td>
<td>3244 ± 108</td>
<td>3237 ± 227</td>
<td>3092 ± 80</td>
</tr>
<tr>
<td>U</td>
<td>197</td>
<td>28</td>
<td>19†</td>
<td>19</td>
<td>43</td>
<td>7†</td>
<td>3</td>
<td>0*</td>
<td>1</td>
<td>0*</td>
</tr>
<tr>
<td>$F_4$ Males</td>
<td>4894 ± 237</td>
<td>4969 ± 142</td>
<td>4622 ± 329</td>
<td>4489 ± 121</td>
<td>4344 ± 663</td>
<td>4115 ± 266</td>
<td>4023 ± 385</td>
<td>4081 ± 345</td>
<td>3748 ± 322</td>
<td>3688 ± 338</td>
</tr>
<tr>
<td>Females</td>
<td>4895 ± 316</td>
<td>4731 ± 270</td>
<td>4549 ± 360</td>
<td>4371 ± 267</td>
<td>4233 ± 141</td>
<td>4240 ± 211</td>
<td>4099 ± 251</td>
<td>4241 ± 177</td>
<td>4364 ± 135</td>
<td>4389 ± 99</td>
</tr>
<tr>
<td>U</td>
<td>247</td>
<td>20</td>
<td>32</td>
<td>15</td>
<td>42</td>
<td>12</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>0**</td>
</tr>
</tbody>
</table>

Note: ' $\leq 10'$ includes ages 8–10, ' $\geq 19'$ includes ages 19–22. Whole number categories (e.g. '11') include all individuals with that whole number age (e.g. 11.00–11.99). Mann-Whitney U-test (two-tailed) represented by the U-value; $^p \leq 0.10$, * $p < 0.05$, ** $p < 0.01$. 

© Informa UK Ltd.
to 3.29) for males and females at each age category. See also Figure 2 for a scatter plot of $F_0$, $D_f$ and $P_f$ values. In order to determine the age when male and female $F_0$ and formant structure diverges, a two-tailed Mann-Whitney U-test was applied to each age category. For $F_0$, males and females were significantly different by age 13 ($U = 1.0$, $p < 0.01$). The difference between male and female formants was not definitive; male and female $P_f$ was marginally significantly different at age 15 ($U = 6$, $p = 0.08$), while $D_f$ did not show a significant difference until the oldest age group ($U = 0$, $p < 0.01$). Low sample sizes at several of the older ages may explain why statistical significance was not achieved for all post-pubertal age categories.

In Figure 2, locally weighted scatter plot smoothing (LOWESS)—a method that models small sections of the data around each point to construct a best-fit curve—was applied to the data in order to explore age-related patterns in $F_0$, $D_f$ and $P_f$ for both males and females. An identifiable pattern is apparent from the LOWESS curves. Males’ $F_0$ decreases relatively slowly until ~12–13 years of age, then rapidly falls until age 17 when the descent levels off. For females, the curve depicts a linear decline that flattens out around age 13.5. Males’ formant structure ($D_f$ and $P_f$) shows a similar pattern: slow descent until age 12, followed by a more rapid drop that tapers off at age ~17.5. Females show a decline of formant structure until age ~15. Due to the relative similarity in age-related change between each formant structure measure, we report only $D_f$ in the following models. Although $P_f$ may show greater sexual dimorphism than $D_f$ (Puts et al. 2012), it uses within-sample means to standardize formant position and, therefore, hampers comparisons across populations and samples.

**Modeling $F_0$ and $D_f$ as a function of age for males and females**

Curve Expert Version 1.5.0 was used to determine a best-fit algorithm for patterns of age-related change in $F_0$ and $D_f$. Goodness-of-fit was assessed using the coefficient of determination ($R^2$). Male voice change and female $D_f$ were best described using a Weibull model, a sigmoid growth function (1), where $a$, $b$, $c$, $d$ and $e$ represent constants (see Table II). Female $F_0$ change was better represented by an MMF function, also a sigmoid function (2).

$$f(x) = a - be^{-cx^d}$$

$$f(x) = \frac{ab + cx^d}{b + x^e}$$

Variance between individuals in rates of growth prevents a perfect fit to the data. Nevertheless, $R^2$ values indicate that the functions fit the pattern of data well (explaining 72% of the variation in male $F_0$ and 59% of the variation in $D_f$). $R^2$ values for females are considerably lower (24% of variance explained for $F_0$ and 46% for $D_f$). Model parameters are listed in Table II. See also Figure 3 for a visual representation of $F_0$ distance and velocity curves.

**Table II. Model parameters for $F_0$ and $D_f$.**

<table>
<thead>
<tr>
<th></th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$d$</th>
<th>SE</th>
<th>$R^2$</th>
<th>APVV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_0^a$</td>
<td>2.43E+02</td>
<td>1.23E+02</td>
<td>4.60E+10</td>
<td>-9.46E+00</td>
<td>29.9</td>
<td>0.72</td>
<td>13.3</td>
</tr>
<tr>
<td>$D_f^a$</td>
<td>1.48E+03</td>
<td>4.50E+02</td>
<td>2.49E+05</td>
<td>-4.70E+00</td>
<td>108</td>
<td>0.59</td>
<td>13.5</td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_0^b$</td>
<td>2.34E+02</td>
<td>2.90E-08</td>
<td>1.18E+03</td>
<td>-9.22E+00</td>
<td>27.9</td>
<td>0.24</td>
<td>n/a</td>
</tr>
<tr>
<td>$D_f^b$</td>
<td>1.48E+03</td>
<td>2.08E+02</td>
<td>4.82E+08</td>
<td>-8.18E+00</td>
<td>86</td>
<td>0.46</td>
<td>n/a</td>
</tr>
<tr>
<td>$F_0^{bc}$</td>
<td>Better condition</td>
<td>2.51E+02</td>
<td>1.66E+02</td>
<td>3.97E+08</td>
<td>-7.64E+00</td>
<td>32.1</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Worse condition</td>
<td>2.38E+02</td>
<td>1.30E+02</td>
<td>5.21E+09</td>
<td>-8.36E+00</td>
<td>22.9</td>
<td>0.82</td>
</tr>
<tr>
<td>$D_f^{bc}$</td>
<td>Better condition</td>
<td>1.51E+03</td>
<td>4.24E+02</td>
<td>2.39E+05</td>
<td>-4.85E+00</td>
<td>116</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>Worse condition</td>
<td>1.46E+03</td>
<td>4.23E+02</td>
<td>3.38E+07</td>
<td>-6.35E+00</td>
<td>101</td>
<td>0.67</td>
</tr>
</tbody>
</table>

*a* Weibull function; *b* MMF function; *c* Males only.

**APVV**, age at peak voice velocity.

Annals of Human Biology
In order to compare Tsimane vocal development with that of a more developed country, we applied the Weibull model to published age-specific \( F_0 \) averages from a North American sample (Huber et al. 1999). Model fit was excellent \((R^2 = 0.99, SE = 6.9)\). Age at minimal velocity is 8.2 (at 248 Hz). Peak \( F_0 \) velocity occurs at age 12.0, at \(-23.1 \, \text{Hz/year}\), when \( F_0 \) is 199 Hz. Maximal acceleration is at age 9.8 (at 241 Hz) and maximal deceleration is at 14.2 (at 153 Hz).

In order to offer a stricter age-matched comparison, we applied a Weibull model to equivalent age categories for the present sample of Tsimane adolescents \((R^2 = 0.93, SE = 23.5)\). All landmarks occur \(\approx 1\) year later in the Tsimane: age at minimal velocity is 9.8 (at 248.5 Hz). Peak \( F_0 \) velocity occurs at age 12.9, at \(-23.1 \, \text{Hz/year}\), when \( F_0 \) is 204 Hz. Maximal acceleration is at age 11.0 (at 240.9 Hz) and maximal deceleration is at 14.8 (at 164.2 Hz).

### Comparison of \( F_0 \) change in Tsimane males in better vs worse condition

A similar analysis was used to compare voice change between males in better vs worse condition. For males in

### Table III. Age and frequency for \( F_0 \) and \( D_f \) growth milestones based on Weibull model.

<table>
<thead>
<tr>
<th></th>
<th>( F_0 )</th>
<th></th>
<th>( D_f )</th>
<th></th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age (years)</td>
<td>Frequency (Hz)</td>
<td>Age (years)</td>
<td>Frequency (Hz)</td>
<td>Age (years)</td>
</tr>
<tr>
<td>Minimum velocity (A)</td>
<td>11.1</td>
<td>243.0</td>
<td>10.1</td>
<td>1485</td>
<td>7.5</td>
</tr>
<tr>
<td>Maximal acceleration (B)</td>
<td>12.0</td>
<td>213.8</td>
<td>11.1</td>
<td>1467</td>
<td>10.1</td>
</tr>
<tr>
<td>Peak velocity (C)</td>
<td>13.3</td>
<td>201.0</td>
<td>13.5</td>
<td>1354</td>
<td>12.4</td>
</tr>
<tr>
<td>Maximal deceleration (D)</td>
<td>14.5</td>
<td>167.1</td>
<td>15.9</td>
<td>1232</td>
<td>14.7</td>
</tr>
<tr>
<td>Length of spur (D–B)</td>
<td>2.5</td>
<td>46.7</td>
<td>4.8</td>
<td>(-235)</td>
<td>4.6</td>
</tr>
<tr>
<td>90% of adult value (E)</td>
<td>16.9</td>
<td>133.2</td>
<td>20.0</td>
<td>1116</td>
<td>17.2</td>
</tr>
<tr>
<td>Length of 90% change (E–A)</td>
<td>5.8</td>
<td>109.8</td>
<td>9.9</td>
<td>(-369)</td>
<td>9.7</td>
</tr>
<tr>
<td>Adult value (F)</td>
<td>23.0</td>
<td>121.0</td>
<td>23.0</td>
<td>1075</td>
<td>23.0</td>
</tr>
</tbody>
</table>

Note: See Figure 3 for a visual guide to milestones A–F.

© Informa UK, Ltd.
better condition, peak $F_0$ velocity occurs at age 13.1, at $-35.3$ Hz/year, when $F_0$ is 198 Hz. For males in worse condition, $F_0$ change occurs over 1 year later at a reduced velocity; peak $F_0$ velocity is at age 14.4, at $-27.8$ Hz/year, when $F_0$ is 195 Hz. Comparing across conditions, for $D_f$, peak velocity occurs at age 12.4 for those in good condition (vs 14.9 for those in poor), at $-59.8$ Hz/year (vs $-65$ Hz/year), when $D_f$ is 1387 Hz (vs 1328 Hz). See Table II for model parameters.

Comparison of voice and height change in this sample

In order to illustrate the magnitude of change in vocal parameters relative to other changes in the pubertal phenotype, we compared $F_0$ and $D_f$ with another widely used pubertal marker—height—using similar metrics. Because of the narrow range of ages in the present sample, the growth curves designed to fit the entire growth period (e.g. JPPS) were not appropriate. The best fit was offered by a MMF model (2), which explained 85% of the variance (SE = 6.8). The first derivative (4) was used to determine velocities.

$$f'(x) = \frac{(bc - ab)dx^d}{x^{2d+1} + 2bx^{d+1} + b^2x} \quad (4)$$

In this model, height increases from 121.6 cm at age 7.5 (at minimum velocity) to 167.1 cm at age 23. Ninety percent of this distance (to 162.6 cm) is completed by age 17.2. Thus, the bulk of change in height (as dictated by this model) occurs over 9.7 years. Peak height velocity is 6.8 cm per year at age 12.5 in this model), which is 2% of height at that age (141 cm). Between the ages of maximal acceleration (age 10.1) and maximal deceleration (age 14.7), height gains 27.2 cm—19.3% of height at peak height velocity (see Table III).

Several comparisons of the magnitude and intensity of change in height compared to $F_0$ and $D_f$ were made for the peri-pubertal ages in this sample. First, voice change occurs over fewer numbers of years: 5.8 for $F_0$ and 9.9 for $D_f$ vs 9.7 for height. Second, peak voice velocity (particularly $F_0$) may be thought of as more temporally compressed: the rate of change in $F_0$ is 16% of $F_0$ at peak velocity (4.2% for $D_f$) vs 2% of height at peak velocity (6.8 cm per year at 141 cm). Third, during the period between maximal acceleration and maximal deceleration, $F_0$ drops slightly more in frequency (23.2% of $F_0$ at peak velocity) in slightly more than half the time (2.5 years) than height does in centimetres (19.3% of height at peak velocity over 4.6 years).

DISCUSSION

To our knowledge this is the first attempt to explore voice change using auxological markers. Modelling the ‘growth’ curve of voice change facilitates quantification of the ages at peak voice velocity, minimum acceleration (i.e. ‘takeoff’ velocity), maximal acceleration and maximal deceleration in $F_0$ and $D_f$. Here we show that these parameters can be used to compare reproductive maturation of boys within and between populations and, therefore, can be utilized as a complement to existing pubertal markers.

Several advantages come with using voice characteristics to characterize the timing of the peri-pubertal shift from growth to reproduction. First, sampling a voice does not require an invasive interview or examination or the presence or participation of a parent or physician. Second, voice recordings are relatively simple and inexpensive to collect, store, transport and analyse. Voice analysis—particularly of $F_0$—is accessible to researchers of varying backgrounds and in different research settings. Third, frequencies are analysed digitally and expressed with an objective number, facilitating cross-sample comparison and limiting inter-researcher error, subjectivity and bias. Fourth, the voice is a socially significant characteristic and is likely to coincide with social and economic changes in males’ roles. Finally, the ‘growth spurt’ in $F_0$ is so pronounced that it can be easily detected and modelled in cross-sectional datasets. The growth spurt in height, by contrast, is sometimes difficult to model in cross-sectional analyses, especially if sample sizes are small or if growth is stunted (Bogin 1999).

Sex differences in $F_0$ and formant structure development during adolescence

Tsimane adolescent males show strikingly different $F_0$ and formant structure development from adolescent females. While female $F_0$ decreases linearly throughout development, male $F_0$ gradually lowers until around age 12–13, when it drops precipitously and significant sex differences emerge. Male formant structure slowly decreases until age 12, followed by a more rapid drop that diverges from females at age 15. Similar patterns of age-related change in male $F_0$ have been observed in several US and European samples (Baken 1987; Hollien et al. 1994; Titze 1994); however, no previous studies have examined voice change in a developing, energetically stressed population where growth and development are often adversely affected (Eveleth and Tanner 1990; Bogin 1999).

Comparison of $F_0$ change in Tsimane and North American males

An age-matched comparison of Tsimane and North American adolescents revealed that Tsimane voice change (as indicated by the age at peak $F_0$ velocity) occurs ~ 1 year later, but at the same $F_0$ and rate of change in $F_0$. A delay in puberty among males in this population accords with observed differences in Tsimane female reproductive timing: Tsimane age at menarche is ~ 13.0 (the current sample) to 13.9 (Walker et al. 2006), whereas the US age at menarche is ~ 12.4–12.6 (Bogin 1999; Herman-Giddens 2006; McDowell et al. 2007). Tsimane youth experience greater energetic stress than typical North American and Western European adolescents, with higher workloads, rates of parasitic infection (Tanner et al. 2009; Vasunilashorn et al. 2010; Blackwell et al. 2011), respiratory and gastrointestinal disease (Gurven et al. 2008) and anaemia (Vasunilashorn et al. 2010). According to life history theory, children and juveniles must trade-off between the competing energy
demands of survivorship and growth. When survival-related energetic demands are high, individuals must delay the onset of pubertal maturation until energy stores are great enough to bear the demands of reproduction (Sterns 1992; Hill 1993; Bogin 1999; Kaplan et al. 2000). Although voices are not directly involved in reproduction, we show that voice change is also delayed in energetically stressed adolescents.

Comparison of \( F_0 \) and \( D_f \) change in Tsimane males in better vs worse condition

Differences in voice characteristics between Tsimane adolescents in relatively better vs worse condition provide a further test of the role of energetic burden in the development of male voices. Tsimane males in better condition had an earlier age at peak \( F_0 \) velocity (age 13.1) compared with those in worse condition (age 14.4). Similarly, males in better condition experienced peak velocity over 2 years earlier than those in poor condition. Variation within the Tsimane exists in many of the factors that affect life history, such as access to clean water, healthcare and high-quality food (Foster et al. 2005; Gurven et al. 2007).

It is notable that, in all groups (North America, better and worse condition Tsimane), peak \( F_0 \) velocity occurred within several Hertz of 200 Hz, despite occurring at variable ages and at disparate velocities in each group. Pedersen et al. (1986) also observed that adolescent \( F_0 \) aggregated into pre- and post-200 Hz groups. Using 200 Hz as a discrete milestone in pubertal development may also be of use to researchers.

Comparison of post-pubertal \( F_0 \) and \( D_f \) in Tsimane males vs other populations

Males in the current sample had a post-pubertal \( F_0 \) of 121 Hz. In comparison, Puts et al. (2012) found a mean \( F_0 \) of 111.4 Hz in a US college population and a mean of 114.7 in a Tanzanian forager population. Other studies have reported similar adult mean \( F_0 \) in US samples (Feinberg et al. 2005; Rendall et al. 2005; Puts et al. 2007). Tsimane adolescent males had a post-pubertal \( D_f \) of 1075 Hz. In a sample of 10 college-aged males aged 20–22, Feinberg et al. (2005) reported an average \( D_f \) of 1005 Hz, while Puts et al. (2012) reported a \( D_f \) of 985 Hz. Listeners attend to both \( F_0 \) and \( D_f \) (Puts et al. 2007; Feinberg et al. 2011) and on both of these dimensions Tsimane males have higher values—and thus higher voices—than US males.

Among energetically-stressed populations, a later age at peak height velocity and a slower rate of growth is associated with shorter stature in adulthood (Bogin 1999; Walker et al. 2006; Godoy et al. 2010). Here we show that Tsimane adolescents have a later age at peak \( F_0 \) velocity as well as a slower rate of \( F_0 \) descent and that this is associated with higher voices among post-pubertal Tsimane males.

Comparison of voice and height change in Tsimane males

In order to illustrate the conspicuous and acute nature of male pubertal voice change, we compared the magnitude and timing of \( F_0 \) and \( D_f \) change with that of height in this sample. Voice frequencies—particularly \( F_0 \)—shift by a greater relative amount in a shorter window of time. One reason for this difference may be that development of the vocal folds is energetically inexpensive. At 17–26 mm (Titze 1994), the vocal folds require little physiological investment compared to general somatic growth (Gurven and Walker 2006). Similarly, \( F_0 \) makes a more dramatic change than \( D_f \) or even height. \( D_f \) is relatively more constrained—by the bony morphology surrounding the vocal tract—than are the small soft-tissue features that determine \( F_0 \).

The fact that voice change—like height—manifests a ‘spurt’ is an interesting phenotypic design. According to Bogin (1999), the late male linear growth spurt has an adaptive function: boys can gain most of the reproductive, economic and social capacities of an adult male while outwardly appearing as a sub-adult. At the end of the adolescent period, when boys are ready to take on the costs associated with participation in the adult world of men, characteristics outwardly identified with manhood, such as height, change rapidly. Because a lower voice is associated with characteristics such as age (Collins 2000; Feinberg et al. 2005), size (Collins 2000; Feinberg et al. 2005) and physical and social dominance (Puts et al. 2006, 2007; Wolff and Puts 2010), the developmental pattern of male voice change may likewise follow this adaptive logic. That is, there may be costs to sounding like a man before one is ready to be a man.

Limitations

Several limitations of the present study should be noted. First, growth of all kinds is better described using longitudinal rather than cross-sectional data. Unfortunately, longitudinal data—especially of the kind needed to accurately measure voice change—are difficult to obtain. Second, sample size for our condition groups and for each age category was relatively small, which increases the chance that extreme values influenced the findings.

Like other measures of reproductive timing, the use of voice characteristics as a pubertal marker also has drawbacks. Collecting high-quality voice samples requires a quiet location with minimal background noise. In addition, voice characteristics may sometimes be affected by sickness, nervousness or other emotions and intentional manipulation. Finally, comparison across samples (i.e. between the Tsimane and North America) is confounded by differences in language. However, care was taken to use words with similar vowels as found in the English language and to use both high frequency (e.g. ‘ee’) and low frequency (e.g. ‘oh’ and ‘ah’) vowel sounds. Finally, it is unclear at this point how our measures align with other pubertal indicators and, for this reason, they may not yet be appropriate for identifying specific aspects of pubertal timing, such as the first onset of puberty, until more research is conducted.

CONCLUSIONS

In the present study, we argue that voice is one of the defining features of adolescent biology and can be used as a marker of the timing and degree of pubertal development.
Past under-utilization of the voice is likely due to the absence of accessible digital tools and processing power, rather than the voice’s inability to appropriately mark pubertal change. We show that traditional auxiological methods can be used to model voice change in adolescents and that this method can be used to compare development both within (i.e. $F_0$ vs $D_0$, voice vs height) and between (i.e. males vs females, better condition vs worse condition, Tsimane vs North America) populations. Future research should explore the factors that affect variance in vocal development within and across populations and in relation to other anthropometric, hormonal and psychological markers of growth and development.

ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the contributions of David Puts. We also thank Tsimane participants and their families, as well as our research assistant, Sofonio Maito Tayo, and his wife, Beronica Canchi Apo, and Bartek Plichta for his help selecting field-appropriate recording equipment and the Tsimane Health and Life History researchers and staff for general support and assistance.

Declaration of interest: This work was supported by a Wenner-Gren dissertation fieldwork grant. The Tsimane Health and Life History project is supported by grants from NSF (BCS-0136761) and NIH/NIA (R01AG024119-01, R56AG024119-06, R01AG024119-07). The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES


Herman-Giddens ME. 2006. Recent data on pubertal milestones in the United States children: the secular trend toward earlier develop-


Hoshi H, Kouchi M. 1981. Secular trend of the age at menarche of Japanese girls with special regard to the secular acceleration of the age at peak height velocity. Hum Bio 53:593–598.


