

**OBJECTIVE:** To collect developmental normative data on the vowel acoustic space and compare against a population with disordered speech and known vocal tract dysmorphology, with the overall goal of establishing anatomic-acoustic correlates through the lifespan for both sexes.

**BACKGROUND:**

- The acoustic vowel quadrilateral area and configuration – demarcated by the point vowels /i u a e/ – define an acoustic space that is reflective of the articulatory working space for vowels.
- Normative data (Flipsen & Lee, 2012; Vorperian & Kent, 2007) inform the study of typical developmental processes and also serve as a reference for the interpretation of data from individuals with speech disorders.
- Deviations in area and shape have been used as indexes of atypical or disordered speech.
- Speech disturbances are common in Down Syndrome (DS, Trisomy 21), due to both anatomic anomalies (vocal tract dysmorphologies) and motor speech disorders (dysarthria, childhood apraxia of speech) and often impair speech intelligibility (Kent & Vorperian, 2013). Studies on vowel acoustics in DS, listed developmentally, variously indicate that:
  - Children aged 3 to 8 years with DS have a smaller ratio of the F2 frequencies for vowels /i/ and /u/ due to their anatomy/maxillary hypoplasia (Moura et al., 2008ab).
  - F1-F2 overlap for different vowels in young speakers may explain listener difficulties in distinguishing vowels (Novak, 1972; and Saz et al., 2009).
  - F2 frequencies for the high vowels are reduced in adolescent participants with DS. However, the difference in F2 between /i/ and /u/ is virtually identical between the DS group and a typically developing control group (Fourakis et al., 2010).
  - There are no differences in formants between adult DS and healthy controls (Moran, 1986).
  - Adult speakers with DS have a smaller acoustic vowel area and a reduced articulatory working space, a reduced range of F2 frequencies for the vowels /i/ and /u/, and reduced F1 frequencies for the low vowels (Bunton & Leddy, 2011).

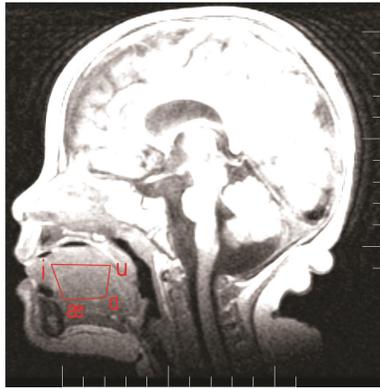


Figure 1. Vowel quadrilateral overlaid on a DS one year old female midsagittal MRI (DS-F178-01-02).

The purpose of this study is to:

- Establish a normative data set of developmental vowel quadrilaterals from typically developing (TD) males and females between the ages 4 to 40 years.
- Present developmental vowel quadrilateral data from males and females with Down Syndrome (DS), ages 4 to 40 years.
- Assess the sensitivity of four metrics to developmental differences between TD and DS.

**METHOD:**

**Participants:**

A total of 256 recordings were obtained from participants between the ages 4 to 40 years. 150 recordings were obtained from 136 typically developing individuals (59 males and 77 females; where some participants returned after a minimum of 1 year); and a total of 106 recordings were obtained from 54 individuals with Down syndrome (57 males and 49 females).

**Speech sample:** Participants repeated 40 monosyllabic words (e.g., “bead” and “bat”) containing the corner vowels, were familiar to younger participants, and had high phonological neighborhood density that reportedly maximize vowel space (Munson & Solomon, 2004). Each vowel was represented in 5 different test words. Stimuli were recorded with a Marantz digital recorder paired with a TOCS+ Platform (speech testing software, Hodge & Daniels, 2007) for randomization. Stimuli were presented visually and aurally.

**Acoustic Analysis.** The waveforms for each word were segmented with the software Praat to select a vowel segment for analysis. The frequencies of the first 4 formants were measured with the software TF32 using FFT spectrograms overlaid with LPC formant tracks. If needed to aid measurements, a spectral slice was generated for the FFT and LPC spectra of a selected interval. Measurements were not made when formant resolution was poor.

The frequency values were used to:

- Create vowel quadrilateral plots, including F1-F2, F1-F3, and F1-F4.

Calculate vowel acoustic space, or F1-F2 planar area.

- Vowel Space Area (VSA)** or the F1-F2 area was computed using following formula for the area of an irregular quadrilateral (Vorperian & Kent, 2007; where  $F_n$  = the formant number for the vowel symbol shown in the virgules):  

$$\text{Area} = 0.5 \times [(F2i \times F1ae + F2ae \times F1a + F2a \times F1u + F2u \times F1i) - (F1i \times F2ae + F1ae \times F2a + F1a \times F2u + F1u \times F2i)]$$

Calculate the following metrics that are reportedly more sensitive than VSA.

- Formant Centralization Ratio (FCR3)** (Sapir, Ramig, Spielman & Fox, 2010).

$$\text{FCR3} = (F2u + F2a + F1i + F1u) \div (F2i + F1a)$$

**Note:** FCR3 is the inverse of vowel articulation index (VAI3).

- Four Vowel Articulation Index (VAI4)** (Karilsson & Doorn, JASA 2012).

$$\text{VAI4} = (F2i + F2ae + F1ae + F1a) \div (F1i + F1u + F2u + F2a)$$

- Vocalic Anatomical Functional Ratio Down Syndrome (VFR)** (Moura, Cunha, Vilarinho, Cunha, Freitas, Palha, Poeschel & Pais-Clemente, 2008a).

$$\text{VFR} = F2i \div F2u$$

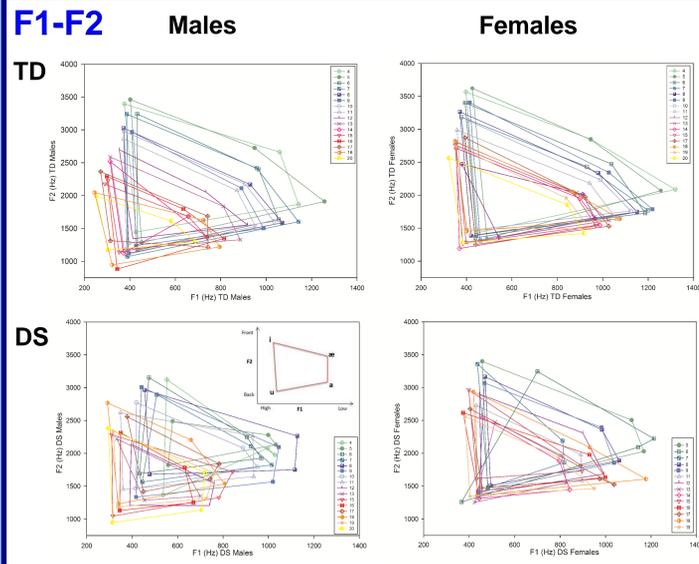


Figure 2 shows F1-F2 plots for males (left) and females (right) in the TD (top panel) and DS (bottom panel) groups.

- TD group:**
  - The vowel quadrilaterals have the shape expected from earlier studies and reflect an orderly developmental decrease in formant frequencies for both sexes.
  - There is a somewhat abrupt shift in both F1 and F2 around age 11-12 years, the approximate age of onset of puberty for both males and females.
  - Sex differences are captured in F1 and F2. F1 differences are reflective of VT length differences as well as possible articulatory differences. For example, large differences in F1 for the low vowels might mean males produce these vowels with a relatively more open jaw position; F2 differences for the high front vowel /i/ might be indicative of oropharyngeal length, width and volume differences (Vorperian & Kent, 2007).
- DS group:**
  - The quadrilaterals are variable in shape, with some individuals having a collapse of the front versus back dimension particularly for the low vowels /a/ and /ae/.
  - The trend of abrupt shift in F1 and F2 around age of 11-12 noted in TD is apparent for DS males but not females.
  - F2 restriction for both low and high vowels in DS appears to reflect limited tongue advancement possibly due to smaller anterior facial skeleton/ oral anatomy.
  - F2 restriction of the low vowels result in collapse of the vowels /a ae/ that may contribute more to reduced speech intelligibility than F2 restriction of the high vowels.

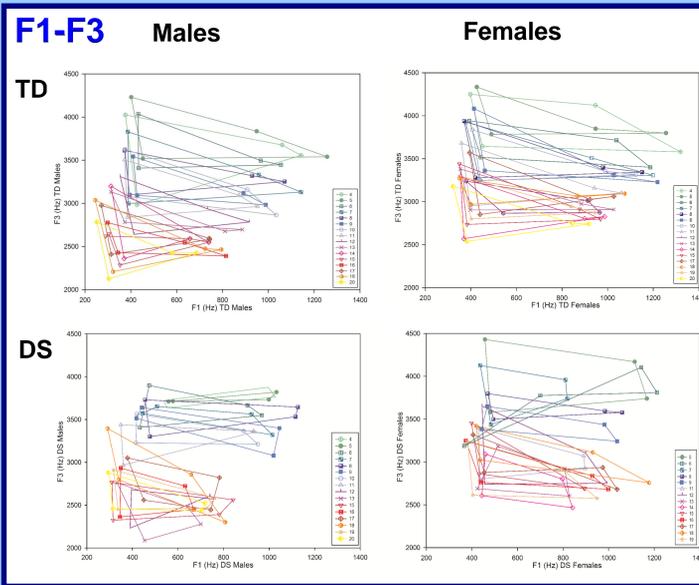


Figure 3 shows F1-F3 plots males (left) and females (right) in the TD (top panel) and DS (bottom panel) groups.

- TD group:**
  - Similar to F1-F2 plots but the acoustic space is more triangular than quadrilateral. As expected, the F1-F3 plots reflect an orderly decrease in formant frequencies with development for both males and females. The F3 values are more dispersed for males than for females.
  - Similar to F1-F2 plots, F1-F3 plots have a somewhat rapid shift in F1 and F3 around ages 11-13 years/puberty in males.
  - The abrupt shifts in F1-F3 data reveal age-group segregation particularly for males.
- DS group:**
  - The vowel acoustic space pattern is irregular during development, and the shape more quadrilateral than triangular (TD).
  - The age-group segregations (of approximately pre-vs-post puberty) are more dominant in the DS group and apparent in both DS male and DS females groups (with females segregating at an earlier age).
  - The F1-F3 plot is suited to developmental studies to detect age- and sex-related differences in vowel acoustics (as opposed to the goal of normalizing these differences in automatic speech recognition).
  - F3 patterns appear to reflect developmental changes in oral cavity anatomy, which are sex-specific and affected by the dysmorphologies in DS.

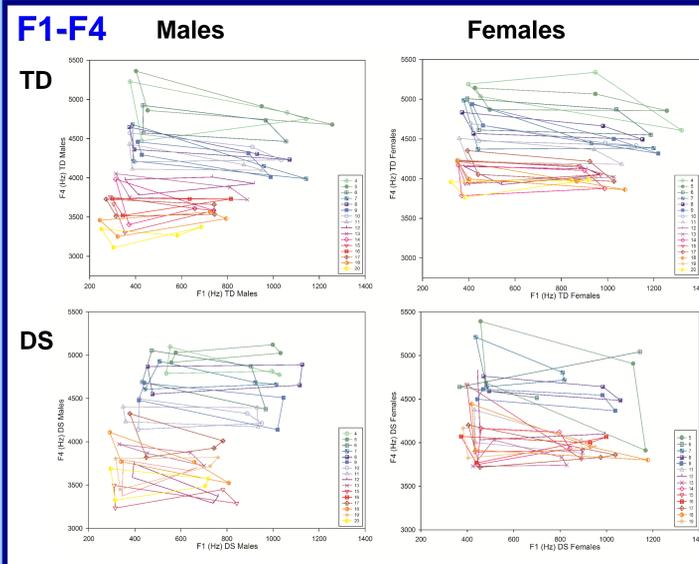
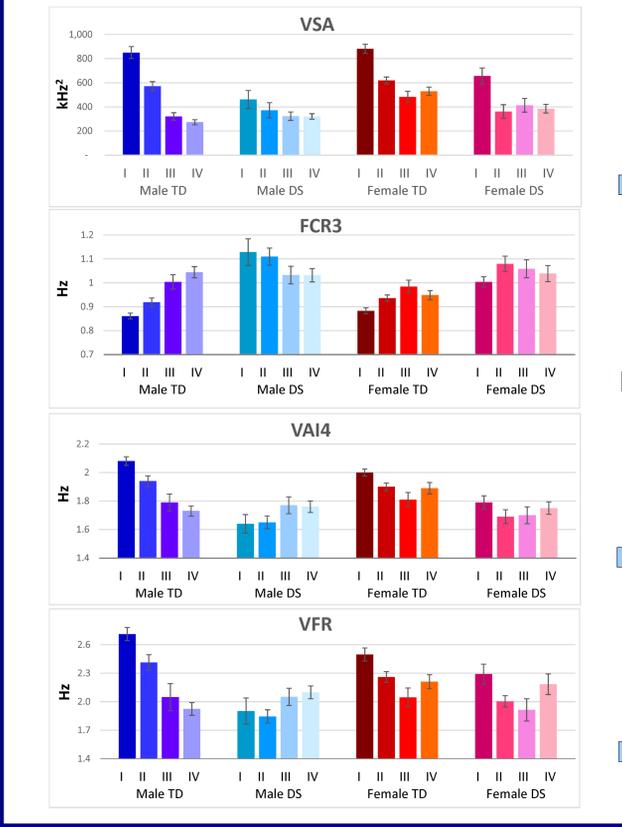


Figure 4 shows F1-F4 plots males (left) and females (right) in the TD (top panel) and DS (bottom panel) groups.

- TD group:**
  - Similar to F1-F3 plots, F4 values are more dispersed for males than females.
  - Similar to F1-F2 and F1-F3 plots, F1-F4 plots have a somewhat abrupt shift in F1 and F4 between ages 11-13 years/puberty in males, and also females.
  - F4 differences are striking between males and females across the developmental range.
- DS group:**
  - The age-group segregations (pre-vs-post puberty) are present but more dominant in males than females.
  - Similar to F1-F3 plots, the F1-F4 plot is suited to developmental studies to detect age- and sex-related differences in vowel acoustics.
  - Sensitivity of F4 to both sex differences and developmental changes particularly during puberty may be reflective of anatomic changes in the hypolaryngeal region during puberty (Takemoto et al. 2006).

Figure 5 shows four metrics (VSA, FCR3, VAI4 & VFR) comparing male (blue shades) and female (red/pink shades) TD (darker color shades) and DS (lighter color shades) speakers divided into four age cohorts: CA of cohort I = 4 to 9 years; II = 10 to 14 years; III = 15 to 19 years; and IV = 20-40 years.



**Metrics (using F1-F2 values):** Based on the following number of recordings per age cohort, calculations reveal:

	TD Male	DS Male	TD Female	DS Female
I	26	13	20	11
II	23	10	25	9
III	7	7	14	10
IV	13	27	22	19

**VSA: A measure of acoustic space.**  
 • TD values, as expected, decrease as age increases reflecting increases in vocal tract length.  
 • DS have smaller vowel area than TD.  
 • DS do not show the systematic drop in area as age increases.

**FCR3: A measure of vowel centralization.**  
 • DS show FCR3 values higher than TD, particularly for young cohorts I & II, implying more vowel centralization.

**VAI4: A measure of vowel contrast.**  
 • DS males and females, particularly young cohorts I and II, show smaller values implying reduced acoustic contrast among vowels.

**VFR: A measure of high vowel /i u/ contrast.**  
 • DS males and females, particularly young cohorts I and II, have smaller values indicating reduced front-back contrast for the high vowels.

**DISCUSSION/CONCLUSIONS:**

- This study provides a normative dataset on the typical development of vowel acoustic space between the ages 4 to 20 years. A unique aspect of this normative reference is that it is based on the same speech material and acoustic analysis methods across the developmental age range studied. Also, it provides data on the first four formants, and is the first study to report developmental F4 data.
- Such normative data permit systematic assessment of developmental changes of the four formants, as well as male-female differences. These data are critical for efforts to establish anatomic-acoustic correlates for diverse speaker groups.
- In addition, these normative data serve as a reference against which disordered speech can be compared using a developmental, sex-appropriate perspective.
- Comparative findings show developmental differences between TD and DS for all formant quadrilaterals and all four metrics.
  - For the younger cohorts I and II, boys and girls with DS have smaller values of VSA (reduced overall vowel area), larger values of FCR (centralization of vowels), smaller values of VAI4 (reduced acoustic contrast among vowels), and smaller values of VFR (reduced contrast for the high vowels).
  - Differences between DS and TD are not as pronounced for the older cohorts III and IV, which may indicate that vowel production in DS improves with development.

- Developmental assessment of anatomic differences between TD and DS can help elucidate aspects of atypical growth observed during development, and its interaction with disturbed motor control.
- Although all current metrics are sensitive to DS speech particularly ages 4 to 15, additional metrics that use higher formants or that better capture low vowel differences can be explored.
- Assessment on the contribution of the different vowels on DS speech intelligibility in the younger DS age group is warranted as the results could have treatment implications.
- The developmental formant frequency patterns for TD and DS speakers may be explained in part by changes in formant-cavity affiliations due to differential growth of the pharyngeal and oral cavities (Martland et al., 1996). Hence, it may be beneficial to consider formant-cavity affiliations in two-tube or three-tube vocal tract models (Apostol, Perrier, & Bailey, 2004).

**REFERENCES:**

1. Apostol, L., Perrier, P., Bailey, G. (2004). *J. Acoust. Soc. Am.*, 115, 337-351.  
 2. Bunton, K., Leddy, M. (2011). *Clin. Linguist. Phon.*, 25, 321-34.  
 3. Desai, S. S. (1997). *Oral Surg., Oral Med., Oral Path., Oral Rad. & Endodontics*, 84, 279-285.  
 4. Flipsen, P., Lee, S. (2012). *Clin. Linguist. Phon.*, 26, 926-33.  
 5. Fourakis, M., Karilsson, H., Tikens, C., Shriberg, L. (2010). *ExLing* 2010.  
 6. Hodge, M., Daniels, J. (2007). *TOCS+ Intelligibility Measures*. Alberta, Canada  
 7. Karilsson, F., van Doorn, J. (2012). *J. Acoust. Soc. Am.*, 132, 2633-2641.  
 8. Kent, R.D., & Vorperian, H.K. (2013). *J. Speech, Lang. Hear. Res.*, 56, 178-210.  
 9. Kumin, L. (2006). *Down's Synd., Res. and Practice*, 10, 10-22.  
 10. Martland, P., Whiteside, S.P., Beet, S.W., Baghai-Ravary, L. (1996). *ICSLP* 1996.  
 11. Moran, M.J. (1986). *J. Comm. Dis.*, 19, 387-94  
 12. Moura, C.P., Andrade, D., Cunha, L.M., Tavares, M.J., Cunha, M.J., Vaz, P., Barros, H., Poeschel, S.M., Clemente, M.P. (2008a). *J. Laryng. & Otolaryg.*, 12, 1318-1324.  
 13. Moura, C. P., Cunha, L. M., Vilarinho, H., Cunha, M. J., Freitas, D., Palha, M., Poeschel, S. M. & Pais-Clemente, M. (2008b). *J. Voice*, 22, 34-42.  
 14. Munson, B. & Pearl Solomon, N. (2004). *J. Speech Lang. Hear. Res.*  
 15. Novak, A. (1972). *Folia Phoniat.*, 24, 182-94.  
 16. Sapir, S., Ramig, L. O., Spielman, J. L., & Fox, C. (2010). *J. Speech Lang. Hear. Res.*, 53, 114-125.  
 17. Saz, O., Simon, J., Rodriguez, W.R., Lleida, E., Vaquero, C. (2009). *EURASIP Journal on Advances in Signal Processing*.  
 18. Takemoto, H., Adachi, S., Kitamura, T., Mokhtari, P., Honda, K. (2006). *J. Acoust. Soc. Am.*, 120, 2228-2238.  
 19. Vorperian, H.K., Kent, R.D. (2007). *J. Speech Lang. Hear. Res.*, 50, 1510-1545.

**ACKNOWLEDGMENTS:** Work was supported by NIH grants # R01-DC 006282 & P30-HD03352. Special thanks for data collection, acoustic analysis and figure preparation to Carlynn Burris, Erin Douglas, Ekatarini Derdemezis, Sara Kurtzweil, Katie Lester, Jen Lewandowski, Erin Nelson, Allison Petska, and Alyssa Wild. Also, many thanks to Simon Lank for assistance with data analysis and figure preparation; and also Ellie Fisher for poster preparation.