

---

# Vowel Acoustic Space Development in Children: A Synthesis of Acoustic and Anatomic Data

Houri K. Vorperian  
Ray D. Kent

University of Wisconsin—Madison

---

**Purpose:** This article integrates published acoustic data on the development of vowel production. Age specific data on formant frequencies are considered in the light of information on the development of the vocal tract (VT) to create an anatomic-acoustic description of the maturation of the vowel acoustic space for English.

**Method:** Literature searches identified 14 studies reporting data on vowel formant frequencies. Data on corner vowels are summarized graphically to show age- and sex- related changes in the area and shape of the traditional vowel quadrilateral.

**Conclusions:** Vowel development is expressed as follows: (a) establishment of a language-appropriate acoustic representation (e.g., F1–F2 quadrilateral or F1–F2–F3 space), (b) gradual reduction in formant frequencies and F1–F2 area with age, (c) reduction in formant-frequency variability, (d) emergence of male–female differences in formant frequency by age 4 years with more apparent differences by 8 years, (e) jumps in formant frequency at ages corresponding to growth spurts of the VT, and (f) a decline of  $f_0$  after age 1 year, with the decline being more rapid during early childhood and adolescence. Questions remain about optimal procedures for VT normalization and the exact relationship between VT growth and formant frequencies. Comments are included on nasalization and vocal fundamental frequency as they relate to the development of vowel production.

**KEY WORDS:** vowels, speech development, formant frequencies, nasalization, vocal fundamental frequency, vocal tract development

---

A half-century ago, Peterson and Barney (1952) published their classic article on vowel formant patterns in men, women, and children, showing that formant frequencies for vowels differ substantially across speakers from different age–sex groupings. Ensuing research has enriched the database on vowel acoustics, and the primary intent of the present article is to consolidate these data into an acoustic portrait of the development of the vowel space from infancy to adulthood in both males and females. The acoustic portrait is supported by information on the anatomic development of the vocal tract, derived primarily from the imaging methods of magnetic resonance imaging (MRI) and computed tomography (CT).

Acoustic methods are a valuable tool in the study of speech development and its disorders, especially because these methods are generally noninvasive, can be readily performed with modern computer systems, and are applicable to a variety of utterance types recorded in laboratory or naturalistic environments. A large number of high-quality recordings of children’s speech are increasingly available for a variety of utterance types, including babbling, early word productions, and

conversation. Therefore, a potentially large database is available for the study of speech development and the adaptation of technologies, such as speech recognition and speech synthesis, to children. As tools for the study of speech development, acoustic studies overcome some of the limitations of perceptual methods, such as biases in phonetic transcription, and they avoid the encumbrances common to many physiologic methods, such as electromyography and movement transduction. To be sure, acoustic analyses have limitations of their own (Kent, 1976; Kent & Read, 2002; Traunmuller & Eriksson, 1997), but technological advances, especially in digital signal processing, enhance the validity and reliability of acoustic analyses of children's speech.

An ultimate goal is the integration of acoustic data with anatomic, physiologic, and perceptual data to produce a comprehensive account of patterns in the development of speech. Such a synthesis would facilitate the interpretation of acoustic data with respect to the other domains of study. This review focuses on the vowel acoustic space in children's speech, interpreted with respect to information on the anatomic development of the vocal tract. This focus was chosen because of the availability of studies that span the developmental period from infancy to adulthood. The primary data under review are the formant frequencies and vocal fundamental frequency associated with vowel production by speakers of various ages and both sexes. The current effort is an update of one part of an earlier article that had a similar goal of summarizing acoustic data on speech development (Kent, 1976). Vowels are important in their own right, but acoustic data on vowels also inform several other topics, including the acoustic cues for consonants (e.g., formant transitions for consonant–vowel or vowel–consonant sequences), speaker normalization (which is usually based on formant frequencies), and prosodic patterns of speech (given that vowels carry a substantial part of prosodic information). In short, vowels are central to an understanding of the acoustic properties of speech. Because vowels appear early in speech development, they are important milestones in the study of speech development. Children achieve a high degree of accuracy in producing nonrhotic vowels by the age of 36 months (Donegan, 2002; Ferguson & Farwell, 1975; Irwin & Wong, 1983; Templin, 1957). This relatively early mastery of vowels relative to many consonants gives vowels a developmental primacy in the establishment of a phonological system.

Acoustic measures of children's speech have a number of applications, including the study of speech development, clinical assessment of speech disorders, technically based interventions for speech disorders, and development of speech recognition systems and speech synthesis systems suitable for children's voices.

However, as considered in more detail later in this article, children's speech presents a number of challenges to acoustic analysis. Acoustic measures of children's speech potentially reflect several developmental processes, including the growth of vocal tract structures (and sex differences in these growth patterns), changes in the relative geometry of the components of the vocal tract, maturation of speech motor control, and convergence on the phonetic patterns of adult speech. These processes are largely concurrent or overlapping, and they may be interactive in their effects. Even though phonetic mastery is typically considered complete by the age of about 8 years, speech development in its finer respects is a protracted process that appears to extend to the late teens in both boys and girls (A. Smith & Goffman, 2004). Interpretation of acoustic data is accordingly challenging, and it would be helpful if the effects of biological factors (such as the growth of the physical apparatus) could be distinguished from factors that reflect phonetic and motor learning.

Contemporary tools allow for a much-improved description of anatomic–acoustic relationships, and these are part of the foundation for a fuller understanding of speech development. Developmental anatomy is discussed separately for the supralaryngeal, laryngeal, and velopharyngeal systems in the Acoustic Correlates of Vocal Tract Length Development, Acoustic Correlates of Laryngeal Development, and Acoustic Correlates of Velopharyngeal Anatomy sections, respectively. These discussions highlight anatomic changes, which provide the biological constraints for speech production, and are critical to the interpretation of developmental acoustic measures.

Chronologic age and speaker sex are the two major determinants of the acoustic properties of speech within a given language. Although chronologic age is not necessarily the preferred independent variable in studies of development or maturation, it is the most frequently reported subject descriptor across studies and, in fact, is typically the only reported index (Kent & Vorperian, 1995). Therefore, chronologic age is the default independent variable used in this developmental description. Combined with speaker sex, chronologic age is the index for studies of maturation and growth.

---

## Acoustic Correlates of Vocal Tract Length Development

The most dramatic effect of growth and development of the vocal tract on vowel production is on formant frequencies, which decrease as the vocal tract lengthens. Vocal tract length in neonates is about 6–8 cm, compared with an average length in adult females of about 15 cm

and in adult males of about 18 cm. We begin this part of the discussion by reviewing recent data on vocal tract anatomy derived primarily from imaging studies.

## Anatomic Considerations

MRI has enabled some of the most comprehensive studies on the growth of the upper airway. This method presents no known biohazard and can be used with subjects of all ages to image both hard and soft tissues in selected planes. Because of the scan time needed for MRI studies and the need to stabilize the head for satisfactory imaging, infants and young children are typically anesthetized for this procedure. The major sources of MRI data on vocal tract maturation are listed in Table 1. The data from these studies provide information on developmental changes in the vocal tract that are of particular importance in accounting for vowel formant-frequency changes with age. The interest is not only on overall length but also how regional growth in the vocal tract (e.g., oral vs. pharyngeal) contributes to vocal tract length (VTL).

Figure 1 shows the measurement of VTL (as defined by Vorperian, Kent, Gentry, & Yandell, 1999) for a 4-year-old male child and a 54-year-old adult male. Vorperian et al. (1999) reported that VTL increased 1.5–2.0 cm during the first 2 years of life and another centimeter between the ages of 25 and 36 months. They also noted that various structures of the vocal tract appear to grow in a synchronized fashion. Fitch and Giedd (1999) observed growth of the pharyngeal region between early childhood and puberty but especially between puberty and adulthood. Arens et al. (2002) concluded that (a) the skeleton of the lower face grows linearly along the sagittal and axial planes for the ages under study and (b) the soft tissues, including tonsils and adenoid, grow proportionately to the skeletal

structures. Vorperian et al. (2005) observed an accelerated growth between birth and 18 months, with no evidence of sexual dimorphism in the growth pattern. They also concluded that the region of the vocal tract (oral–anterior vs. pharyngeal–posterior) and orientation (horizontal vs. vertical) determine the developmental growth pattern. Although the pharyngeal–posterior structures account for vocal tract lengthening throughout development, growth of oral–anterior structures is particularly prominent during the first 18 months of life. These anatomic changes are pertinent not only to ontogeny but also to evolutionary proposals that attempt to account for the unique two-tube vocal tract configuration in humans (Nishimura, Mikami, Suzuki, & Matsuzawa, 2006).

## Anatomic–Acoustic Relationships

Certainly, a basic principle in relating anatomic change to acoustic correlates is that the length of the vocal tract determines the overall pattern of formant frequencies. As children mature, their vocal tracts lengthen, and their formant frequencies decrease. However, the actual pattern of formant-frequency change as a function of age may not be simple because the growth of the vocal tract is not just a matter of uniform lengthening. Particularly in males, the vocal tract has disproportionate growth in the pharyngeal region compared with the oral region. Fant (1975) suggested the following relationships between cavity length and formant frequencies:

$$\text{Pharyngeal cavity length} = 35300/2 \times F2;$$

$$\text{Oral cavity length} = 35300/2 \times F3.$$

Thus, according to Fant, the pharyngeal cavity length is affiliated with the second formant, and the oral cavity length is affiliated with the third formant. Childers and Wu's (1991) findings are supportive of the second formant affiliation whereby they report F2 to be a slightly better recognizer of gender than fundamental frequency in adults. Perry, Ohde, and Ashmead (2001), on the other hand, reported that at age 4 years (the youngest age they studied), F3 was lower for boys than for girls, with small differences in F1 and F2. Whiteside (2001) also noted that even before puberty, there is a considerable tonotopic distance between the F3 values of males and females. Interestingly, Lieberman, McCarthy, Hiiemae, and Palmer's (2001) findings have shown that although there are no apparent sex differences in the distance between the posterior pharyngeal wall to the lips (supralaryngeal vocal tract–horizontal), the oropharyngeal portion of the supralaryngeal vocal tract–horizontal (i.e., oropharyngeal width—the distance from the posterior pharyngeal wall to the posterior margin of oral cavity) is slightly larger in males between the ages of 1.75 and

**Table 1.** Major sources of magnetic resonance imaging data on the developing vocal tract.

Study	n	Age
Fitch and Giedd (1999)	129	2–25 years
Vorperian et al. (1999)	2 <sup>a</sup>	Birth–3 years, 9 months
Vorperian (2000)	20 <sup>b</sup>	Birth–6 years, 9 months
Arens et al. (2002)	92	1–11 years
Vorperian et al. (2005)	37 <sup>c</sup>	Birth–6 years, 9 months

Note. For each published study, the table shows the number of imaging cases/studies and case age at time of imaging.

<sup>a</sup>Longitudinal data were used for the children. <sup>b</sup>Some children were studied longitudinally. <sup>c</sup>Composed of 25 children (some of whom were studied longitudinally) and 12 adults.

**Figure 1.** The measurement of vocal tract length (VTL) defined as the curvilinear distance along the midline of the tract starting at the thyroid notch to the intersection with a line drawn tangentially to the lips. Left panel is the midsagittal magnetic resonance imaging (MRI) of a pediatric male subject at age 4 years, 4 months, with VTL measuring 11.28 cm. Right panel is the midsagittal MRI of an adult male subject at age 54 years, 2 months, with VTL measuring 15.87 cm.



4.75 years. An alternative conclusion presented by Martland, Whiteside, Beet, and Baghai-Ravary (1996) is that there is a transposition of the F2 and F3 parameters owing to differential growth of the pharyngeal and oral cavities during development, such that for children younger than 2 years, F3 is related primarily to the pharyngeal cavity. Thus, formant-cavity affiliation may not be limited to cavity length only but also cavity width. This idea is further supported by Robb, Chen, and Gilbert (1997), who reported that formant frequencies remain fairly stable during the first 2 years of life while there are documented increases in VTL (Vorperian et al., 2005). Also, there are reports that speaker sex identification prior to 10–12 years is based on the resonance characteristics of the vowels (Perry et al., 2001) while there are no significant differences in VTL (Fitch & Giedd, 1999) and no significant differences in fundamental frequency (see *Vocal F<sub>0</sub> Data From Database Sources* subsection of the *Acoustic Correlates of Laryngeal Development* section). Therefore, it may be more accurate to characterize the nonuniform growth of the vocal tract as nonuniform growth of length, width, and subsequently volume. Whiteside (2001) also noted that in addition to nonuniform sex differences in the VTL (pharyngeal cavity length, oral cavity length, and total VTL), there is the need to investigate sex differences in vocal tract volume. Ultimately, such information can be integrated in articulatory models, such as the variable linear articulatory model developed by Maeda (1979, 1990), and applied developmentally, as done by Menard, Schwartz, and Boe (2004). The use of articulatory models that account for both the nonuniform growth of length and width of the vocal tract should help advance the understanding of exchanges and interplay of formant-cavity affiliations.

## Formant-Frequency Patterns Across Development

The sources of formant-frequency data reviewed in this article are from 14 of the 21 studies listed in Appendix A.

## Value of Formant Descriptions

Formant descriptions are a low-dimensional description of vowels, although formants are not necessarily superior to other acoustic representations for various purposes, including perceptual representations and speaker normalization (De Wet et al., 2004; Molis, 2005; Zahorian & Jagharghi, 1993). One advantage of a formant specification is the systematic relationship between formant pattern and vowel articulation (which is to say, the acoustic-to-articulatory conversion). The classic F1–F2 formant plot depicts a fundamental articulatory–acoustic relationship in which the F1 and F2 frequencies are related principally to tongue height and advancement, respectively. Alternatively, the F2–F1 difference can be interpreted as tongue advancement–retraction. Data on vowel formant frequencies in children have been reported in a sufficient number of studies, particularly ages 3 years and up, to yield a satisfactory composite data set to summarize developmental patterns (see Appendix B). Data on F1 and F2 are most abundant, but a few studies also have reported data on F3. Given the cavity affiliation issues noted above in the *Anatomic–Acoustic Relationships* section, an F1–F2–F3 description is desirable for a reasonably complete description of vowel development because F3 complements F1 and F2 information, particularly with respect to speaker normalization and the identification of rhotic

vowels. The most common methods used to estimate formant frequencies in children are spectrograms, automated routines (such as linear predictive coding), or both of these used together. To our knowledge, there has been much less use of other techniques, such as acoustic impedance spectrometry (Epps, Dowd, Smith, & Wolfe, 1997), cepstral analyses (Fort & Manfredi, 1998), or acoustic reflection technology (Xue & Hao, 2003, 2006).

## Estimation Error

It is always important to assess measurement error in determining the precision of formant frequencies, but this error takes on even greater importance in studies that use variability of formant frequencies as an index of maturation, with the usual hypothesis being that formant-frequency variability (and presumably, therefore, articulatory variability) diminishes with age. That is, the error in formant-frequency estimation can be confounded with the variability associated with intraspeaker imprecision in achieving articulatory-acoustic targets. Distinguishing measurement error from maturation-related variability is one of the challenges of acoustic analysis. From an analytic point of view, the error of formant-frequency estimation is related to  $f_0$  because higher  $f_0$  values result in a larger spacing of harmonics. Generally, the closer the spacing of the harmonics, the better defined are the peaks of the vowel spectrum.

Age-related variability of formant-frequency pattern in vowel production has been determined in several studies. One of the earliest systematic developmental studies was a cross-sectional investigation by Eguchi and Hirsh (1969), who showed essentially continuous decreases in the variability of both F1 and F2 from 3 to 11 years of age. However, Nittrouer (1993) reported that F1 variability was minimal by the age of 3 years, whereas F2 variability continued to decrease after that age. She interpreted this result to mean that precision of jaw movement (which affects especially F1) was achieved relatively early. The relative maturation of motor control over different oral structures is not entirely clear. Children's jaw movements are less variable than lip movements (Green, Moore, & Reilly, 2002; Walsh & Smith, 2002), but it has been reported that jaw and lip movements have parallel decreases in variability with maturation (Walsh & Smith, 2002).

Aside from the above noted challenge of using variability of formant frequencies to distinguish between measurement error and articulatory variability as an index of maturation, there is the additional complication of separating intra- versus interspeaker (within- vs. between-speaker) sources of variance. Furthermore, there is the difficulty of interpreting the origin of interspeaker sources of variance because it seems that, concurrent with periods of decreased articulatory variability,

there is a decrease in the anatomic growth rate of various vocal tract structures, particularly during early childhood (Vorperian, 2000; Vorperian et al., 2005). Thus, elucidation of the sources of variability in speech development rests on the availability of multiple types of data (including acoustic, anatomic, and movement data).

## Data Sources

Searches were made of major bibliographic databases (PubMed, PsycLIT) and selected journal indexes (*The Journal of the Acoustical Society of America* and *Journal of Speech, Language, and Hearing Research*) to identify studies reporting data on vowel formant frequencies. The search terms were as follows: vowels, formants, formant frequency, speech acoustics, and speech development. As noted above, 21 source studies of formant-frequency data are listed in Appendix A, along with descriptions of the speech samples used and their analysis method. The studies were further examined to determine their suitability for inclusion. Of the 21 source studies, 14 candidate studies were identified according to the following criteria: (a) studies reported quantitative data on developing (child) or mature (adult) speakers of English, (b) developmental data were reported for more than one single age group, (c) data were reported for at least three, but preferably four, of the corner vowels, (d) group studies were preferred over single-subject reports, and (e) quantitative data were reported for at least the first two formant frequencies (F1 and F2). The next step was to calculate average formant values per vowel per age group to graphically summarize the data to depict developmental relationships. Particular emphasis was given to the classic vowel quadrilateral because of the general availability of data for the corner vowels and the utility of the quadrilateral in defining the overall vowel acoustic space and articulatory-acoustic correlates establishing this space.

As can be seen in column 6 of Appendix A, the formant-frequency data used in this study were from speakers of various geographic regions. Thus, the age and gender comparisons described in this article are confounded with dialect variation. Ideally dialect should be taken into consideration in the interpretation of data from any particular study and, more specifically, the place of birth and childhood residence for the characteristics of low vowels and high back vowels (Clopper, Pisoni, & de Jong, 2005). One reason why dialectal influence was difficult to control is because the formant-frequency data used in this study were published over an interval of nearly 5 decades, and dialects shift over time. Another confounder was that most of the studies included in the present analysis did not ascertain that the subjects did in fact have the dialect typical of speakers from a given geographic region.

## Graphical Analysis

Vowel quadrilaterals were created by first identifying the subset of studies from the 14 candidate studies appropriate to each plot (male, female, child). Male plots present data for males from childhood (in which sex is specified at 4 years of age) through adulthood. Similarly, female plots present data for females from childhood (in which sex is specified at 4 years of age) through adulthood. Child plots present data only for subjects younger than 12 years of age (with an average of male and female values when sex is specified). The F1 and F2 values (and F3 values when available) reported in 14 of the 21 studies listed in Appendix A were used to generate the age–sex-indexed average quadrilaterals in Figures 2–7. The last two columns in Appendix A specify what is included from each study in the various plots (male, female, and child). Appendix B lists the studies and data sources per age group. The corners of the quadrilaterals are simple means derived for each of the four corner vowels /i/, /u/, /æ/, and /a/ from all appropriate studies for a given age group. Study–age combinations with missing vowels were deleted (details are available in Appendix A). In this way, the data for each age (or age–sex) group summarize the published data.

The F1–F2 vowel quadrilaterals are shown in Figure 2 (males ages 4 years through adulthood), Figure 3 (females ages 4 years through adulthood), and Figure 4 (children ages 9 months to 11 years). The legend in Figures 2–4 includes data on the areas of the vowel quadrilaterals (vowel acoustic space size) at the different ages. F1–F2 planar area was computed with the following formula for the area of an irregular quadrilateral:

$$\begin{aligned} \text{Area} = & 0.5 \times [ (/i/F2 \times /æ/F1 + /æ/F2 \times /a/F1 \\ & + /a/F2 \times /u/F1 + /u/F2 \times /i/F1) \\ & - (/i/F1 \times /æ/F2 + /æ/F1 \times /a/F2 \\ & + /a/F1 \times /u/F2 + /u/F1 \times /i/F2) ], \end{aligned}$$

where  $F_n$  = the formant number for the vowel symbol shown in the virgules; for example, /i/ F2 is the second formant for vowel /i/.

The prediction from standard acoustic theory is that vowel formant frequencies decrease as the vocal tract lengthens with age. This prediction is supported by the data in Figures 2, 3, and 4. Although the data on vowel quadrilateral area data are somewhat variable across studies, a general decline in quadrilateral size is evident during development (see Figure 8). The variability in the results is not surprising, given the multiples sources of formant-frequency data used to construct the composite graphs. Data on vowel acoustic space size in normal vowel development are a useful reference for the study of

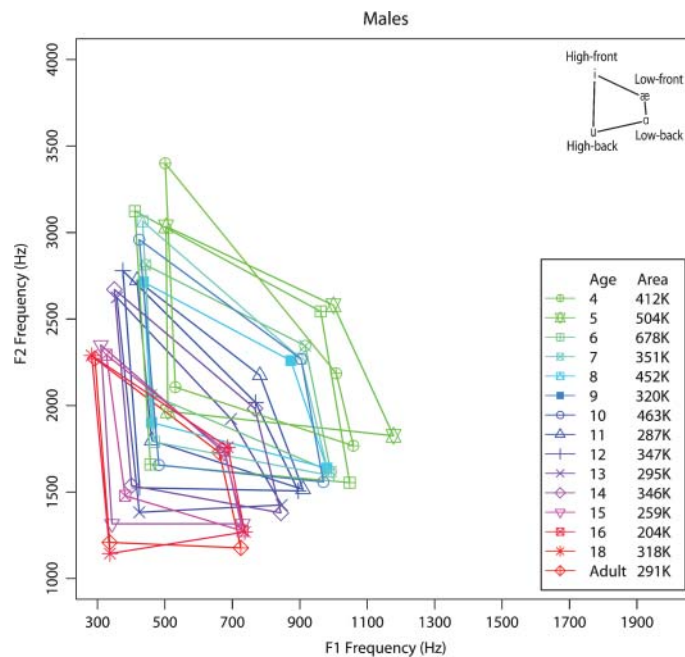
children with dysarthria, deafness, and various developmental disorders (Higgins & Hodge, 2001; Kent, Netsell, Osberger, & Hustedde, 1987; Liu, Tsao, & Kuhl, 2005; Moura et al., in press; Rvachew, Slawinski, Williams, & Green, 1996; Schenk, Baumgartner, & Hamzavi, 2003). Unusually small areas are correlated with reduced intelligibility in children and with possible risk for speech disorder. Furthermore, vowel-specific formant-frequency differences may have value in characterizing the vocal tract features of particular syndromes (Moura et al., in press). Therefore, development of vowel space size is one index of the capacity for intelligible speech, and normative data can help in the acoustic interpretations of unintelligible speech.

The F1–F3 data generally take the form of a quadrilateral, but there are some exceptions to this geometry, such as a reversal of the configuration for the back vowels (see discussion of figures in the next paragraph). The composite F1–F3 data are shown in Figure 5 (males 4 years through adulthood), Figure 6 (females 4 years through adulthood), and Figure 7 (children 8.5 months to 11 years). A fairly regular age-related pattern can be seen in the F1–F3 plots, but there is a conspicuous decrease in F3 between the ages of 1 and 3 years. Also, the F1–F3 quadrilaterals have a greater developmental dispersion or separation, that is, there is less overlap of the quadrilaterals than the F1–F2 quadrilateral patterns, particularly for males. This may indicate that the F1–F3 analyses are more sensitive to age and possibly to speaker sex, as noted above (see the *Anatomic–Acoustic Relationships* section).

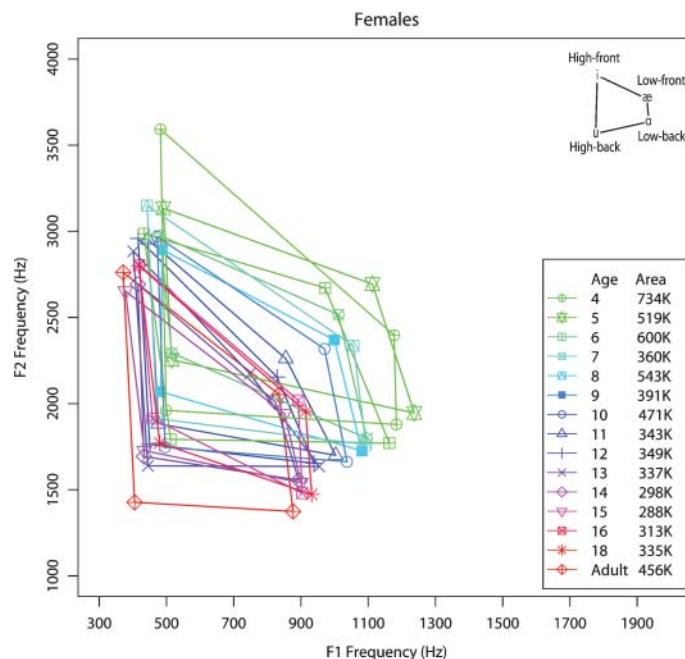
Figures 9 and 10 show F1–F2 and F1–F3 measurements from the study of Perry et al. (2001), who reported data for boys and girls at the ages of 4, 8, 12, and 16 years. These data are particularly instructive regarding age–sex differences in formant patterns because they allow an inspection of age–sex related changes in the vowel quadrilateral. A sex difference in the acoustic space begins to emerge even in the data for 4-year-olds, especially for the low vowels in which the F1 values are about 150–200 Hz lower for males than for females. This difference becomes more pronounced with age, such that progressively less overlap is noted in the vowel quadrilaterals for the two sexes. By the age of 16 years, the quadrilaterals do not overlap. An additional potentially interesting feature is that there is a sex difference in F1 frequency for low vowels across all age groups, with males having lower F1 values.

Average F1–F2 data for adults from eight studies are shown in Figures 11 and 12. These illustrations are collections from relatively large  $N$  studies of speakers of English. These data for adults are shown here for comparison purposes in the study of speech development and to show the variation in formant-frequency data for

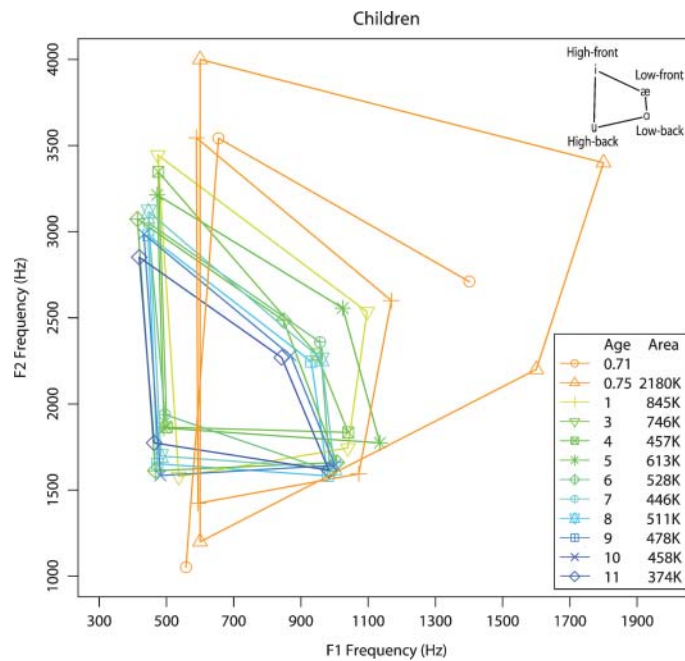
**Figure 2.** Average F1–F2 acoustic space for males (ages 4 years through adulthood) from 12 of the 21 studies listed in Appendix A. Plotted data are averages across studies, at a given age, for the four corner vowels. Separate vowel quadrilaterals formed from these averages are shown at each age, and the area of each of the vowels spaces is given in the inset to the figure. The column before last in Appendix A indicates the studies that had data available for formant average calculations and lists the specific ages for which data were available for averaging.



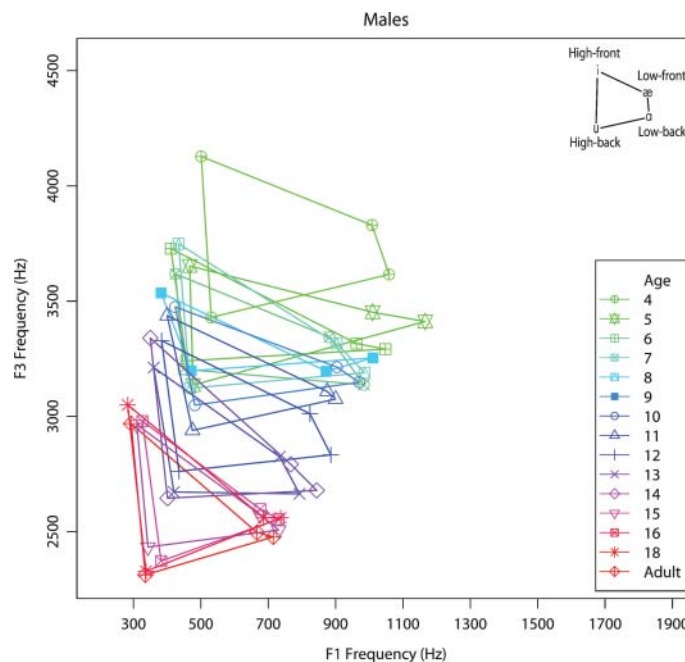
**Figure 3.** Average F1–F2 acoustic space for females (ages 4 years through adulthood) from 11 of the 21 studies listed in Appendix A. Plotted data are averages across studies, at a given age, for the four corner vowels. Separate vowel quadrilaterals formed from these averages are shown at each age, and the area of each of the vowels spaces is given in the inset to the figure. The column before last in Appendix A indicates the studies that had data available for formant average calculations and lists the specific ages for which data were available for averaging.



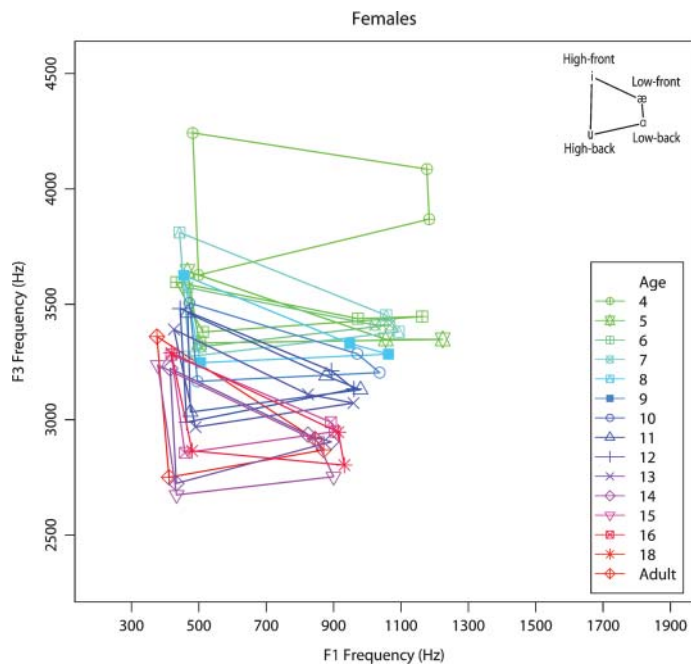
**Figure 4.** Average F1–F2 acoustic space for children (ages 8 months to 11 years) from 11 of the 21 studies listed in Appendix A. Plotted data are averages across studies, at a given age, for the four corner vowels. Separate vowel quadrilaterals formed from these averages are shown at each age, and the area of each of the vowels spaces is given in the inset to the figure. The column before last in Appendix A indicates the studies that had data available for formant average calculations and lists the specific ages for which data were available for averaging.



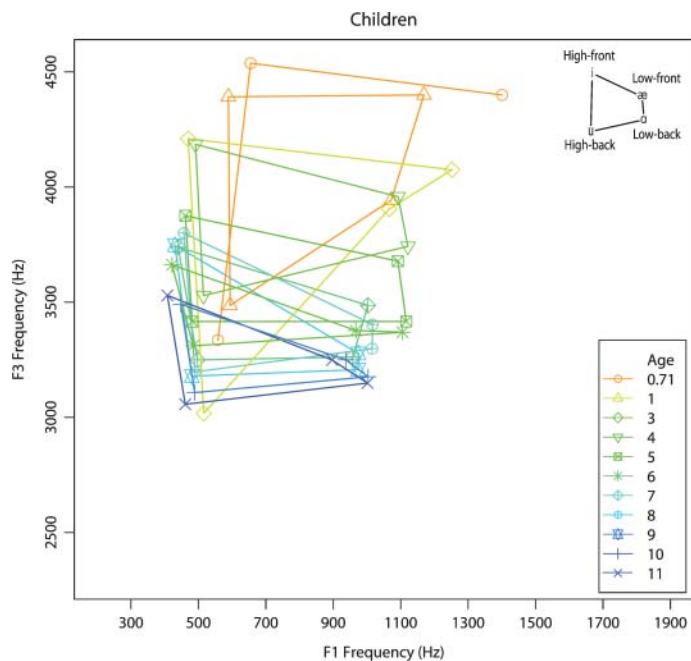
**Figure 5.** Average F1–F3 acoustic space for males (ages 4 years through adulthood) from 10 of the 21 studies listed in Appendix A. Plotted data are averages across studies, at a given age, for the four corner vowels. Separate vowel quadrilaterals formed from these averages are shown at each age. The last column in Appendix A indicates the studies that had F3 data available for formant average calculations and lists the specific ages for which data were available for averaging.



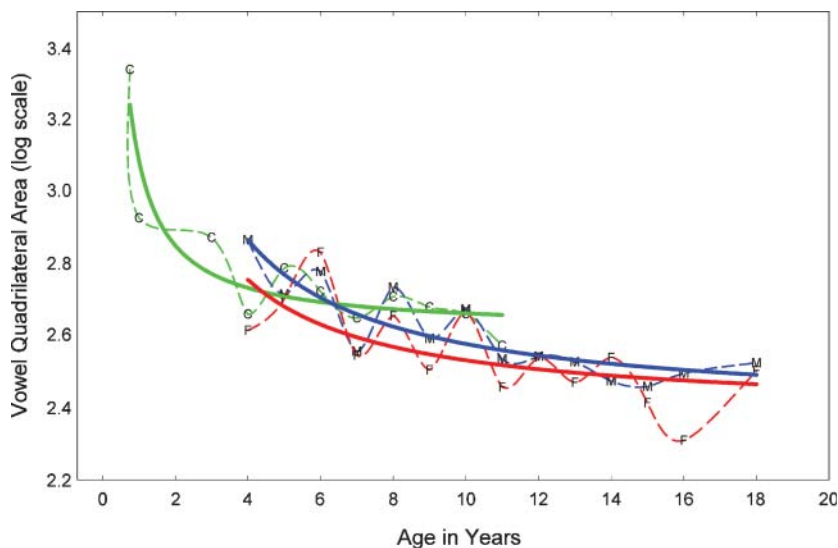
**Figure 6.** Average F1–F3 acoustic space for females (ages 4 years through adulthood) from 9 of the 21 studies listed in Appendix A. Plotted data are averages across studies, at a given age, for the four corner vowels. Separate vowel quadrilaterals formed from these averages are shown at each age. The last column in Appendix A indicates the studies that had F3 data available for formant average calculations and lists the specific ages for which data were available for averaging.



**Figure 7.** Average F1–F3 acoustic space for children (ages 8.5 months to 11 years) from 8 of the 21 studies listed in Appendix A. Plotted data are averages across studies, at a given age, for the four corner vowels. Separate vowel quadrilaterals formed from these averages are shown at each age. The last column in Appendix A indicates the studies that had F3 data available for formant average calculations and lists the specific ages for which data were available for averaging.



**Figure 8.** Vowel quadrilateral areas (log scale of Hz squared/1000) of the average F1–F2 acoustic space per age group for males (M–blue), females (F–red), and children (C–green; as displayed in Figures 2, 3, and 4), with three distinct cubic polynomial regression fits per plot (males, females, and children).



speakers in whom maturational processes are presumed to be complete. Although the phonetic context of the words from which the vowels were analyzed can affect the vowel acoustic space (Munson & Solomon, 2004), it is also likely that formant frequencies may continue to change somewhat during adulthood, apparently because of continuing growth of the human cranial skeleton (Israel, 1968, 1973). Data in support of this possibility have been reported in several studies that demonstrate increases in size of the various craniofacial structures well into late adulthood (Endres, Bambach, & Flosser, 1971; Linville & Rens, 2001; Rastatter, McGuire, Kalinowski, & Stuart, 1997; Scukanec, Petrosino, & Squibb, 1991; Xue & Hao, 2003).

### Acoustic Evidence of Growth Spurts

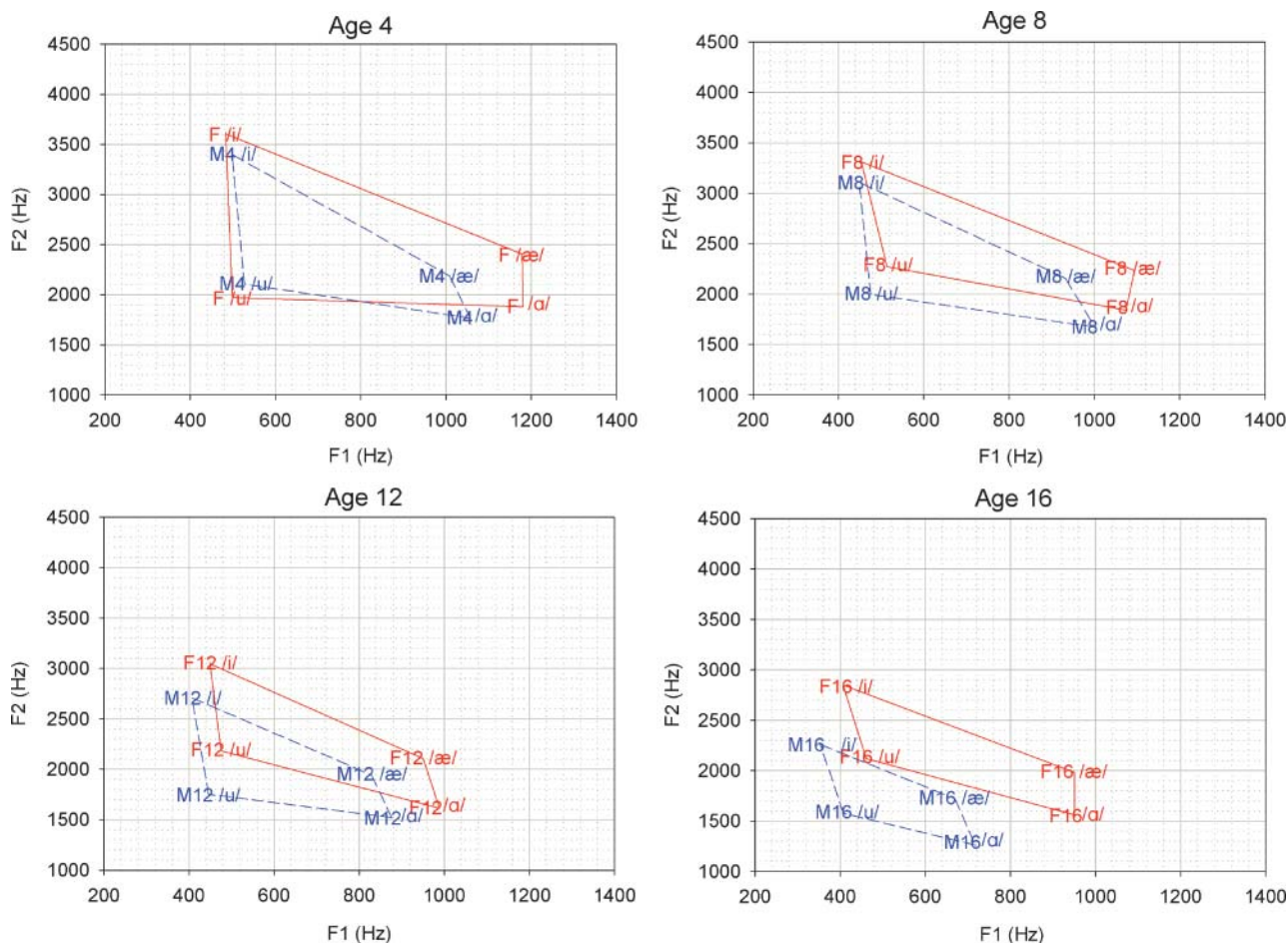
An important developmental question is whether anatomic growth spurts at certain ages can be identified from acoustic data. Because of the large variance in the data from published studies, it is difficult to answer this question with confidence. However, some tentative conclusions can be offered, beginning with the period of infancy.

*An exception to the standard prediction from acoustic theory.* Robb et al. (1997) concluded from a cross-sectional study of 20 children that average F1 and F2 frequencies were essentially stable over the period from 4 to 25 months of age, but they did observe a significant decrease in the average bandwidths for both F1 and F2. Bandwidth data have been rarely reported in

developmental studies. Variations in bandwidth speak to changes in absorption of sound by vocal tract tissues or possibly to subtle changes in nasal resonance. In a study of four children over the developmental period of 15–36 months of age, Gilbert, Robb, and Chen (1997) noted essentially constant F1 and F2 frequencies before 24 months (and, by interpretation, little change in VTL) but significant decreases in both formant frequencies between 24 and 36 months (and presumably a lengthening of the vocal tract). To the contrary, MRI data show rapid increases in VTL in the first 2 years (Vorperian, 2000; Vorperian et al., 1999, 2005). Possibly, the formant-frequency results cannot be explained solely by anatomic changes of increases in VTL. For example, it may be necessary also to examine changes in pharyngeal length and width in relation to formant frequency and bandwidth changes. The study by Robb et al. appears to be the only source of developmental data on formant bandwidth. As mentioned earlier, the reduction in formant bandwidth observed in this study could be the result of reduced nasalization and/or a change in the biomechanical properties of the tissues of the vocal tract or volumetric changes in the pharyngeal region. Nasalization is further discussed later in this article (see the Acoustic Correlates of Velopharyngeal Anatomy section), and it appears that changes in velopharyngeal function may very well account for reductions in formant bandwidth in the first 2 years of life.

*Jumps in vowel acoustic space.* An interesting observation based on Figures 2–7 is that across the age

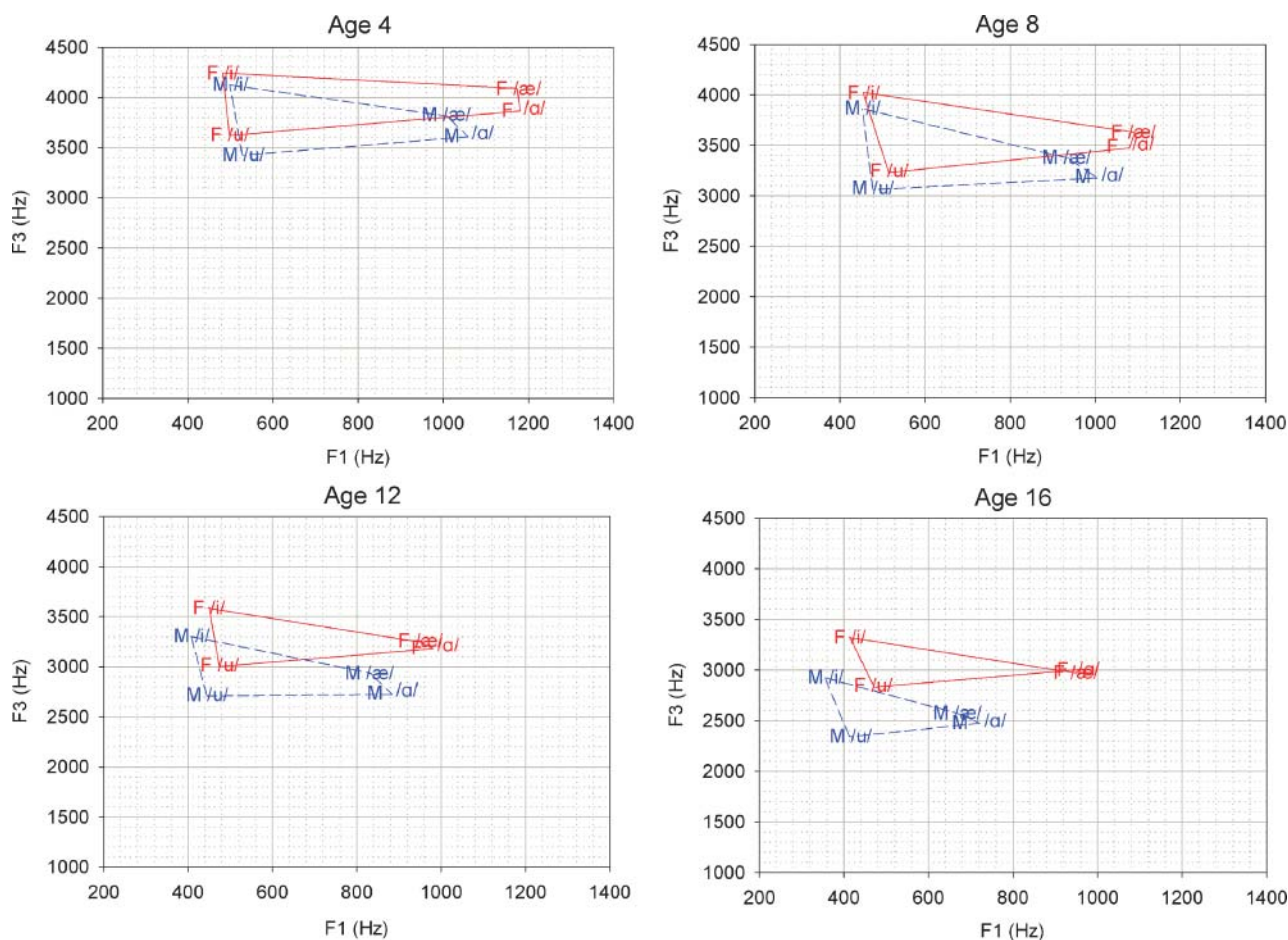
**Figure 9.** F1–F2 data from the study by Perry, Ohde, and Ashmead (2001), who reported data for boys and girls at the ages of 4, 8, 12, and 16 years. The corner vowels are the averages from 20 subjects per age group (10 males, dashed blue lines; 10 females, solid red lines), with five repetitions per vowel. Sex differences in the acoustic space are evident in the data for the 4-year-olds. This difference increases with age, and there is progressively less overlap in the vowel quadrilaterals for the two sexes. An interesting observation is that there is a sex difference in F1 for the low vowels across all age groups, with males having lower F1 values.



increments plotted, there are notable jumps or skips in the F1–F2 and F1–F3 vowel acoustic data between particular age groups. That is, changes in formant frequencies are nonlinear with respect to chronological age. Two types of jumps can be noted, an overall jump in vowel acoustic space and a limited jump in the low vowel region of the vowel acoustic space. In Figure 2, summary of male acoustic data, there is a noticeable overall jump in the F1–F2 vowel acoustic space between the ages of 14 and 15 years when abrupt drops in F1 and F2 formant frequencies can be noted for all corner vowels. For example, between the ages 14 and 15 years, the first and second formant frequencies for the low-back vowel /æ/ drop about 100 Hz and 250 Hz, respectively. In Figure 4, summary of child acoustic data, there is a noticeable overall jump in F1–F2 vowel acoustic space between the ages 1 and 4 years, that is, an abrupt change in the first and second formant frequencies for all corner vowels.

Similar overall jumps in the F1–F3 acoustic space can be noted at similar ages in Figure 5 (males) and Figure 7 (children). It is reasonable to relate these jumps in vowel acoustic space to the primary descent of the larynx during infancy and the secondary descent of the larynx during adolescence, particularly in males (Fitch & Giedd, 1999). Abrupt increases in VTL cause abrupt decreases in formant frequencies and hence a jump in vowel acoustic space. A slightly different jump in vowel acoustic space is apparent in Figure 3, summary of female acoustic data, in which between the ages 10 and 12 years, there is a jump in the F1–F2 acoustic space that is limited to the low vowels. For example, for the low-front vowel /æ/, the average of both F1 and F2 values drops about 150 Hz. A similar jump in the F1–F2 acoustic space that is again limited to the low vowels can also be noted in Figure 4, summary of child acoustic data, between the ages 5 and 6 years. Identical jumps can be noted in the F1–F3

**Figure 10.** F1–F3 data from the study by Perry, Ohde, and Ashmead (2001), who reported data for boys and girls at the ages of 4, 8, 12, and 16 years. The corner vowels are the averages from 20 subjects per age group (10 males, dashed blue lines; 10 females, solid red lines), with five repetitions per vowel. Sex differences in the acoustic space are evident at 4 years of age and increase with age. The overlap in the vowel quadrilaterals for the two sexes decreases, and by age 16 years there is no overlap. Note a sex difference in F1 for the low vowels across all age groups, with males having lower F1 values.

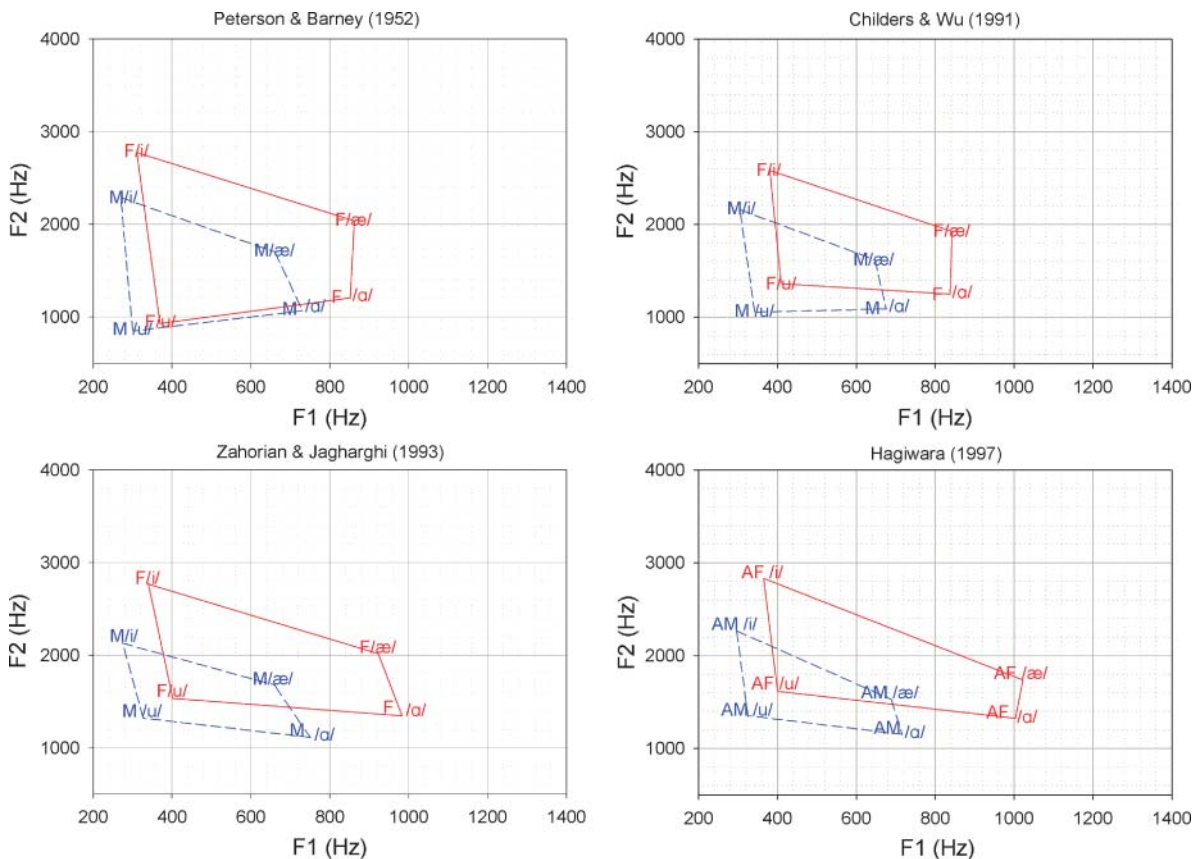


acoustic space in Figure 6 (females) and Figure 7 (children). Figure 13 is a composite display of the males', females', and children's average F1, F2, and F3 values for all four corner vowels that were used to create the vowel quadrilaterals in Figures 2–7. Abrupt and concurrent changes in all formant frequencies at various ages are apparent, including the ones described above.

Although quantitative anatomic data on the developing vocal tract are limited, it is well known that the growth of the vocal tract is nonuniform, for example, the ratio of the pharyngeal (posterior) region to oral (anterior) region of the vocal tract is larger for adult males compared with adult females and children (Fant, 1975). Thus, it is reasonable to further postulate that differences in jumps in the vowel acoustic space—overall versus limited to low vowel region—are related to differences in the anterior–oral versus posterior–pharyngeal

regions of the vocal tract. On the basis of Figures 3 and 4, it appears that such differences become evident between ages 5 and 7 years and are well established between the ages 10 and 12 years. Indeed, Lieberman et al. (2001), using a longitudinal series of radiographs, determined that the ratio of pharynx height to oral cavity length decreased significantly between birth and 6–8 years. They also observed that certain aspects of vocal tract shape changed markedly during the 1st postnatal year and during adolescence. Additional sex-specific quantitative anatomic data on the developmental changes in the length of the anterior–oral or horizontal vocal tract region and height of the posterior–pharyngeal or vertical vocal tract region, especially in conjunction to data on the width, area, or volume in the pharyngeal region, would be of value because there is physical evidence that by age 12 years, boys have a larger neck circumference (Bennett, 1981; Perry et al., 2001).

**Figure 11.** Average F1–F2 data from four studies with adult values (male-blue, female-red). The studies are listed in Appendix A. Note the similarities and also the variation in the formant-frequency data for speakers in whom the maturational processes are presumably complete.

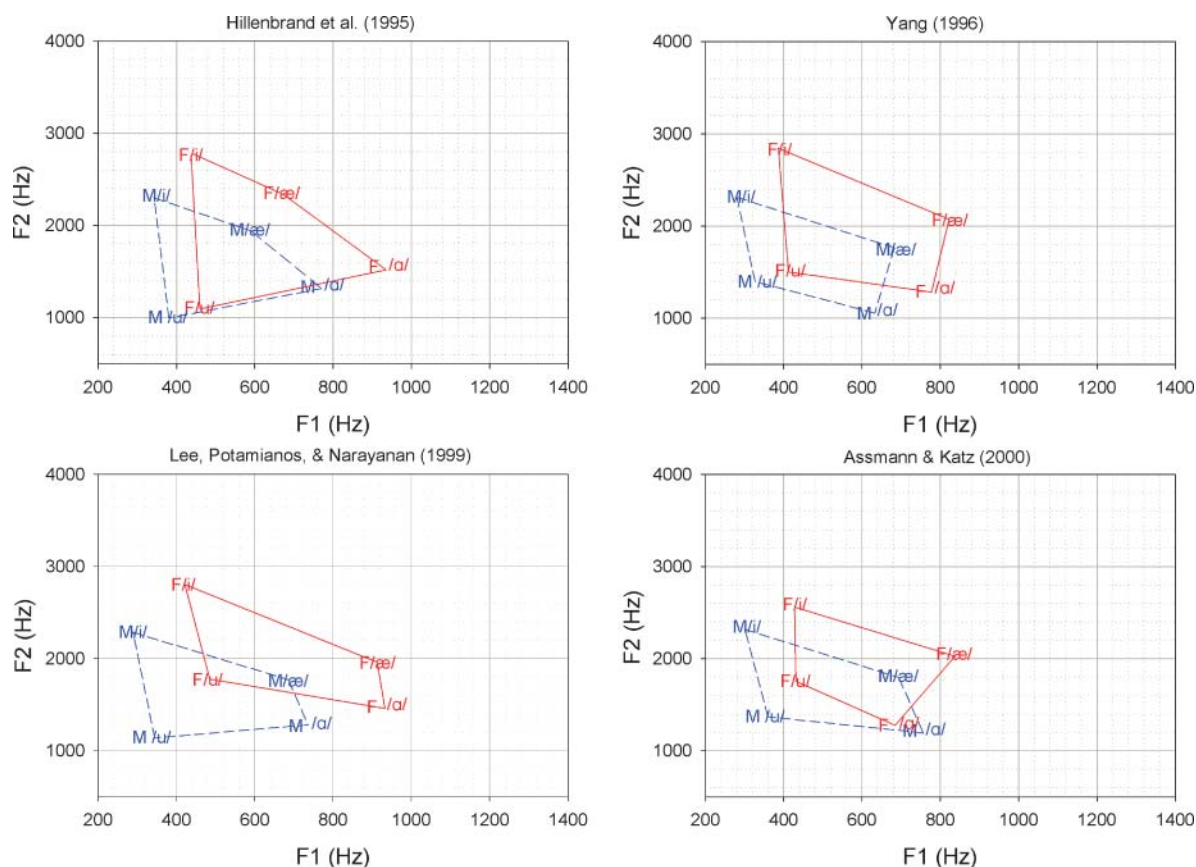


*Variability roots.* As reviewed in the Acoustic Correlates of Vocal Tract Length Development section, formant-frequency variability is a measure that is typically used to assess inter- and intraspeaker articulatory variability. An interesting observation can be seen in Figures 14, 15, 16, and 17 comparing the average F1, F2, and F3 values for the different vowels during the course of development. In general, there is minimal variability in F1 for the high vowels but increased variability in F2 particularly for the high-back vowel /u/ across the entire developmental age range. Nittrouer (1993) concluded that the emergence of mature gestural patterns is not uniform. Similarly, the growth of the pharyngeal versus oral regions of the vocal tract is not uniform (Fant, 1975). The increased variability of F2 for the high-back vowel /u/ across the entire developmental age range may indicate that variability is rooted in several factors, including the influence of dialect, articulatory variability, and variability due to the nonuniform anatomic growth of the vocal tract, particularly in the posterior pharyngeal region, which—as noted above in the *Anatomic-Acoustic Relationships* section—is typically affiliated with the second formant.

## Sex Differences

At some point in development, males and females have vocal tracts that differ in length and shape. It appears from the composite acoustic data in Figures 2–7, and Figures 14–17, that sexual dimorphism of the vocal tract emerges by the age of 4 years, and the differences become more apparent by age 7 or 8 years when boys have consistently lower formant frequencies than girls across all vowels (Bennett, 1981; Busby & Plant, 1995; Lee, Potamianos, & Narayanan, 1999; Perry et al., 2001; Whiteside & Hodgson, 2000). Additional acoustic differences become more apparent after age 12 years when discrete male–female differences in  $f_0$  are evident (see Figure 18) and as significant differences in VTL emerge (Fitch and Giedd, 1999). Thus, the acoustic data converge on the conclusion that sex differences in speech acoustics begin in early childhood well before puberty. The identification of speaker sex before age 12 years must be predominantly due to differences in the resonator–vocal tract but not its length. Childers and Wu (1991) have noted that F2 is a better recognizer of gender than  $f_0$ . As seen in Figures 2–7 and 14–17, the pattern of F1, F2, and particularly F3 dispersions—differences

**Figure 12.** Average F1–F2 data from four additional studies with adult values. The studies are listed in Appendix A. Note the similarities and also the variation in the formant-frequency data for speakers in whom the maturational processes are presumably complete.



for the different vowels is not consistent for males versus females. Thus, to determine the anatomic correlate for such developmental acoustic differences in F1, F2, and F3, it is necessary to assess empirically the changes in pharyngeal length–height and width–area–volume during the course of development and to compare the pharyngeal dimensions with those of the oral portion of the vocal tract.

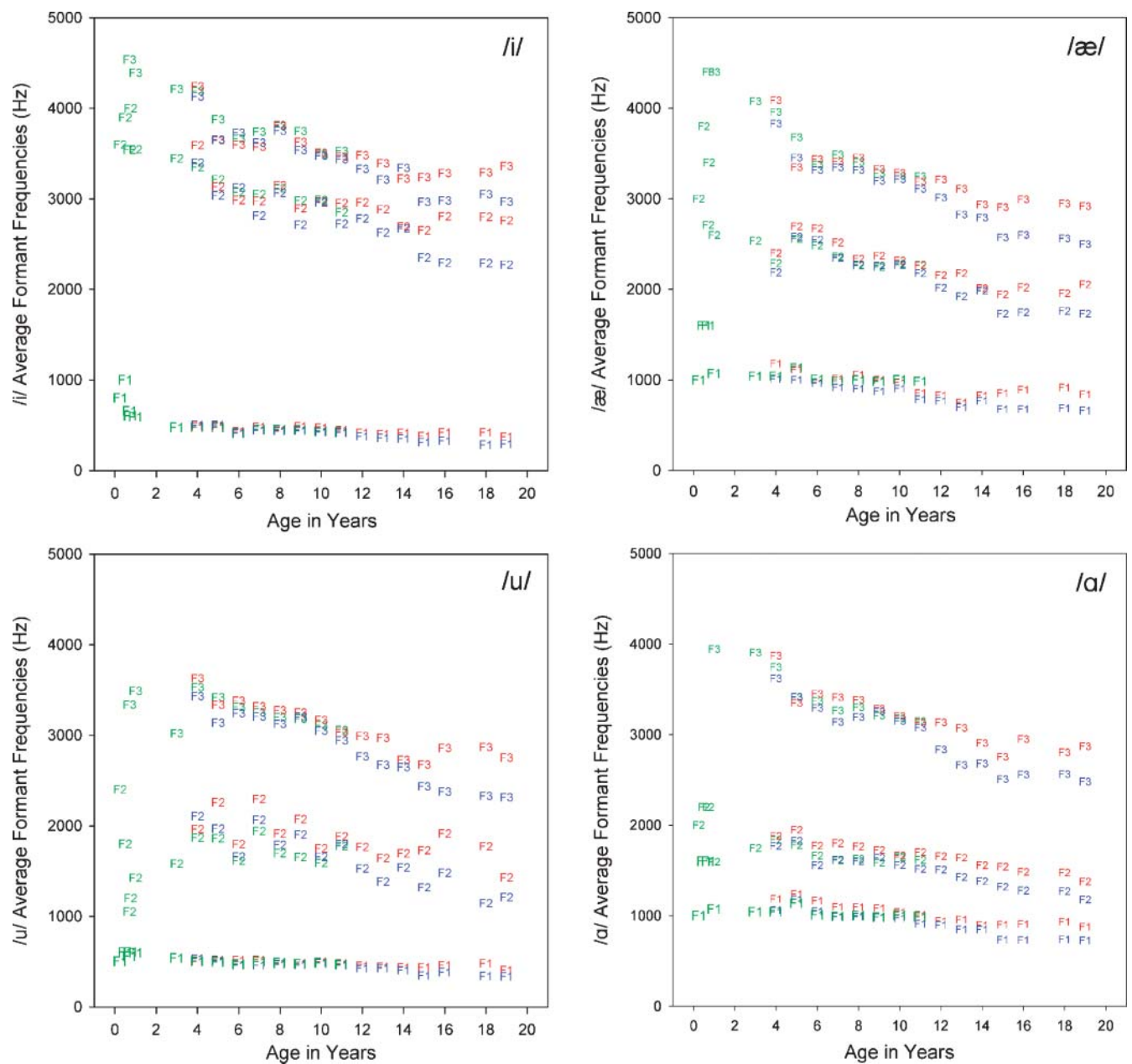
### Vocal Tract Normalization

The problem of vocal tract normalization (also known as speaker normalization) is a longstanding issue in acoustic phonetics and, more recently, in speech technologies, such as automatic speech recognition (Fant, 1975; Martland et al., 1996). The formant frequency differences summarized in this article motivate the need for scaling factors that normalize for age–sex differences in the acoustic properties of speech. Normalization for VTL is complicated by an apparent sex or gender difference in the articulation of low vowels and by idiolectal–dialectal differences in vowel production. As noted in Figures 2, 3, 5, and 7, the largest sex differences

in vowel formant frequencies occur for F1 for the low vowels /æ/ and /a/ and for F2 for the vowel /i/. These differences in vowel formant frequencies may reflect some articulatory differences between boys and girls in addition to differences in VTL and more specifically differences in anterior–oral versus posterior–pharyngeal portions of the vocal tract that affect formant-cavity affiliations. For example, the large difference in F1 frequencies for the low vowels might mean that boys produce these vowels with a relatively more open jaw position, and differences in F2 frequencies for the high-front vowel may be indicative of sex differences in oropharyngeal length, width, and volume, as noted in the *Anatomic–Acoustic Relationships* section.

It is not entirely clear whether a uniform scaling factor suffices to normalize vowel formant frequencies for both boys and girls (Fant, 1975; Kent, 1976; Lee et al., 1999; Martland et al., 1996; Whiteside & Hodgson, 2000). Lee et al. (1999) observed a linear change in formant frequencies for males between the ages of 11 and 15 years and concluded that their data were consistent with the hypothesis of uniform axial growth. However, in a reanalysis of the data of Lee et al., Whiteside (2001) concluded

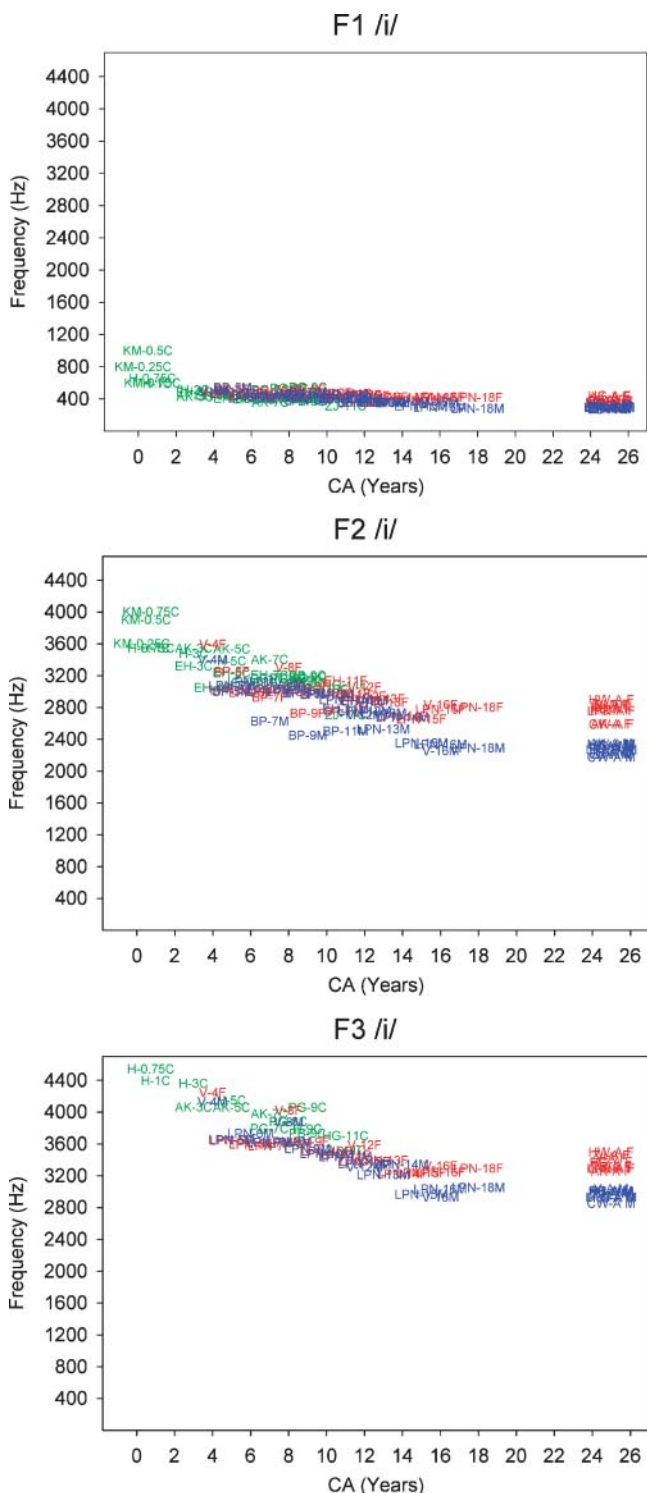
**Figure 13.** A composite display of the males', females', and children's average F1, F2, and F3 values across studies for the four corner vowels across age. The formant averages are the same as those displayed in the vowel quadrilaterals in Figures 2–7.



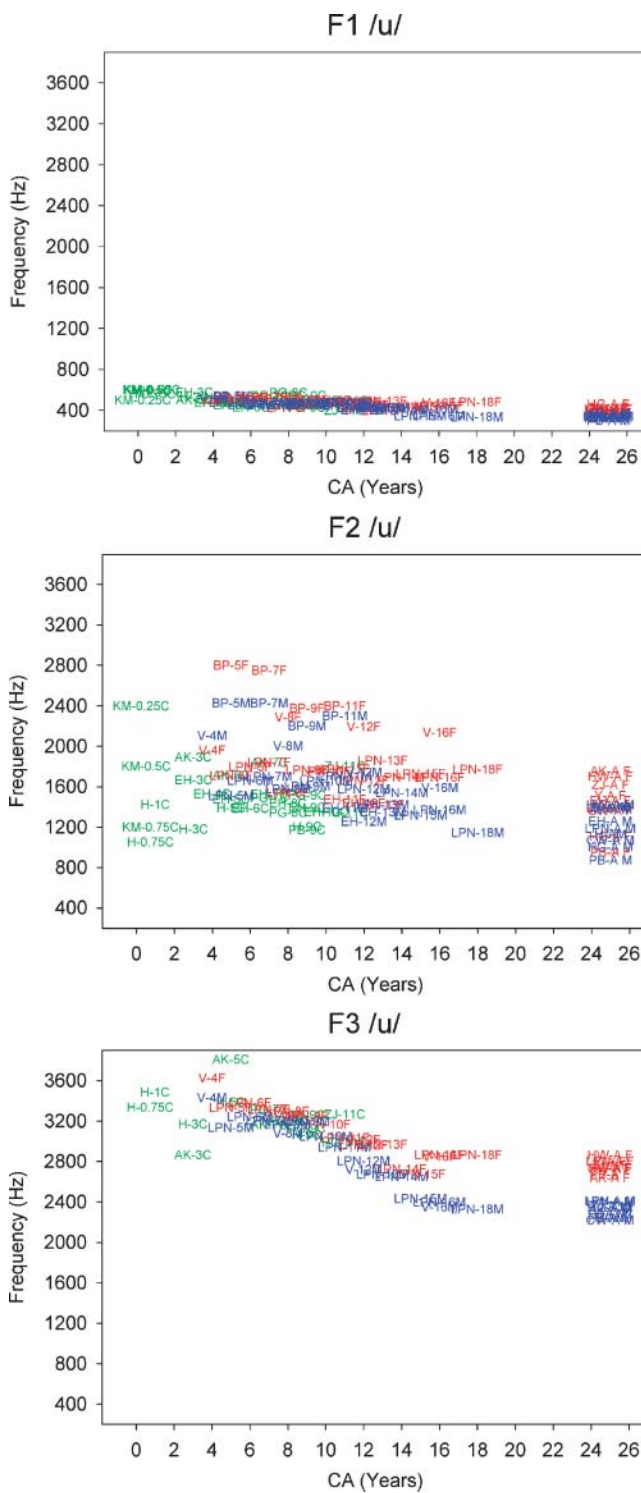
that there was a nonlinear increase in the tonotopic distance between male and female data for vowel formant frequencies. Similarly, White (1999) concluded that vowel-dependent formant frequency differences between boys and girls indicated nonuniform differences in the dimensions of male and female vocal tracts. White also noted that these sex differences were not consistent with data for adult vowels. White's data for 29 children who were 11 years of age showed that formant frequencies were higher for speech than for singing and also were higher for girls than for boys.

The findings in this article indicate that although the prediction from standard acoustic theory holds that vowel formant frequencies decrease as the vocal tract lengthens with age, such decreases are not necessarily linear with chronological age, as noted by the jumps or skips in formant frequencies at particular ages for each sex. Also, although there are noted developmental and sex differences in the anterior–oral versus the posterior–pharyngeal portions of the vocal tract, those differences are not limited to length of the cavities, particularly the posterior–pharyngeal cavity but also cavity width and

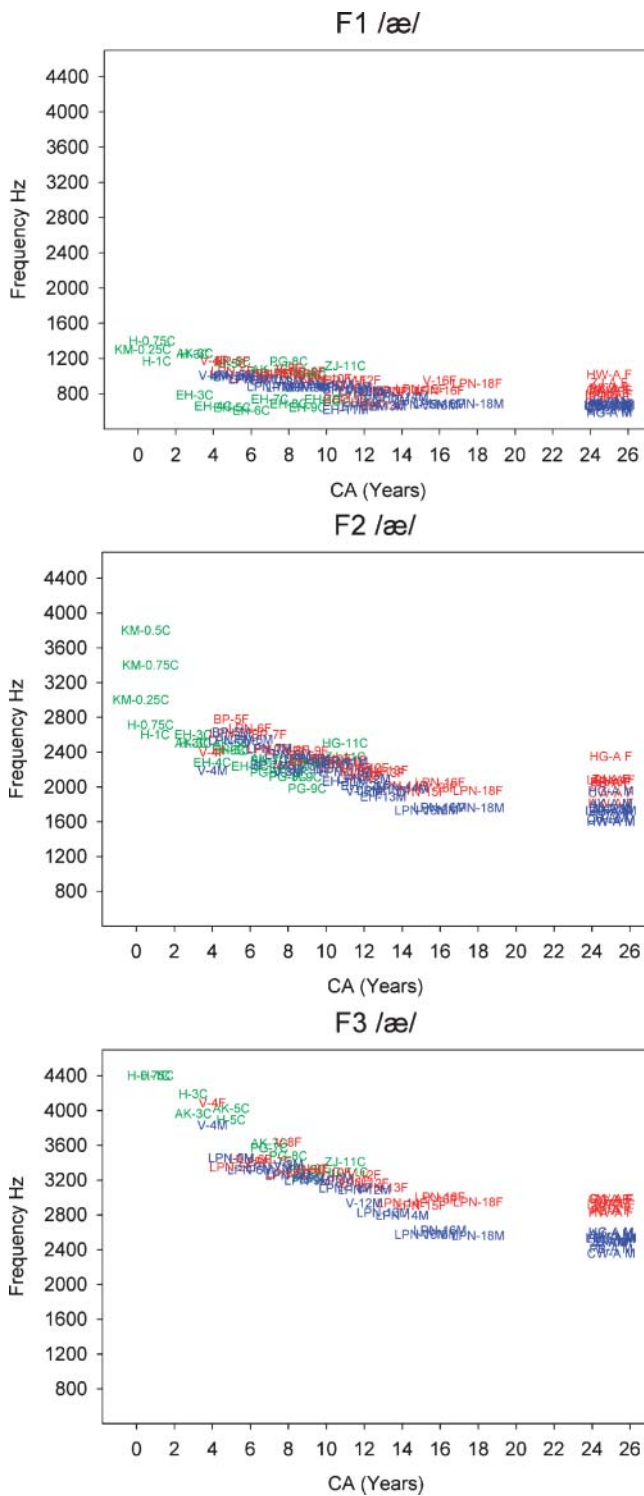
**Figure 14.** Average F1, F2, and F3 values per study by chronological age (CA) for the vowel /i/ during the course of development for the three groups: children (green), females (red), and males (blue).



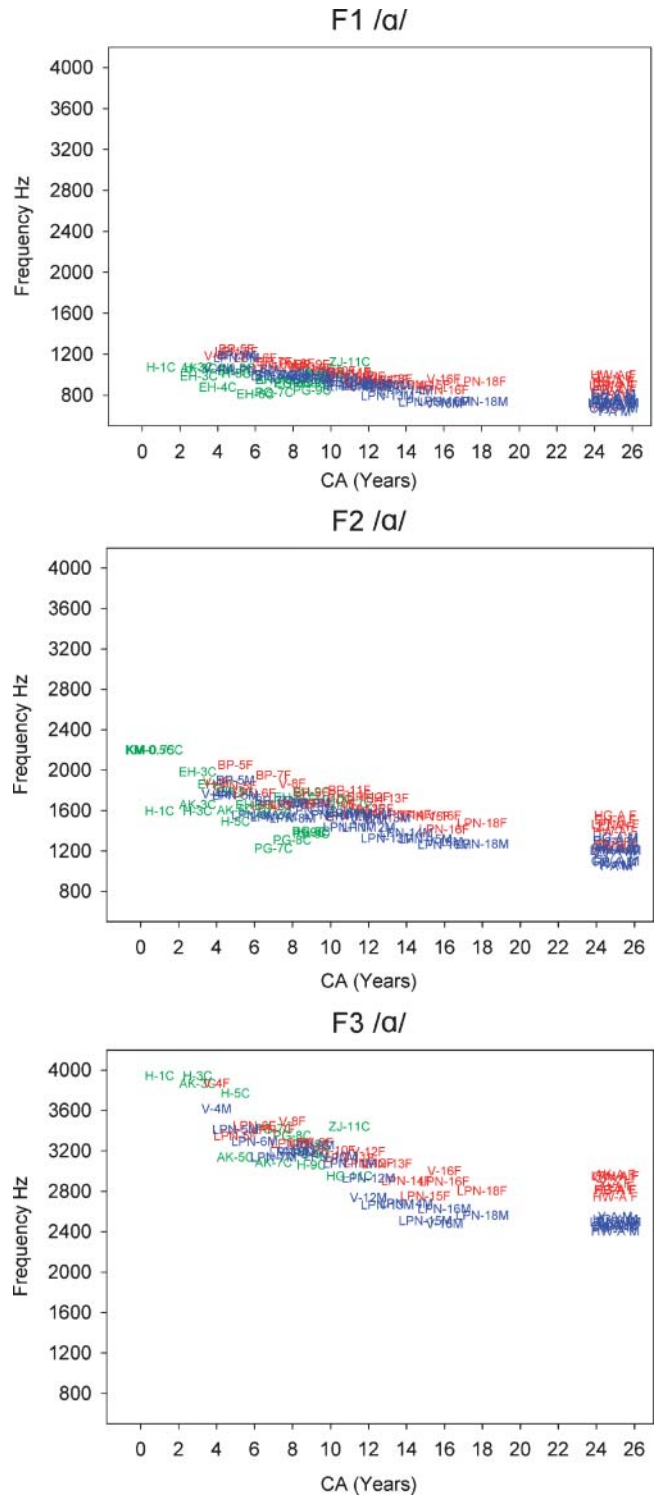
**Figure 15.** Average F1, F2, and F3 values per study by chronological age (CA) for the vowel /u/ during the course of development for the three groups: children (green), females (red), and males (blue).



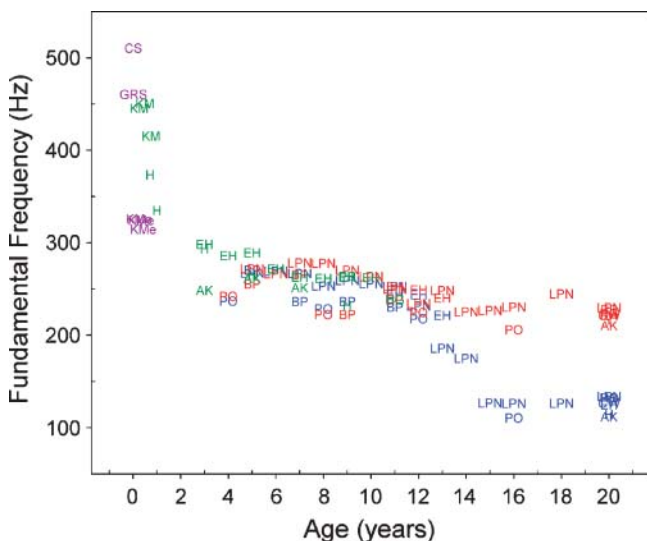
**Figure 16.** Average F1, F2, and F3 values per study by chronological age (CA) for the vowel /æ/ during the course of development for the three groups: children (green), females (red), and males (blue).



**Figure 17.** Average F1, F2, and F3 values per study by chronological age (CA) for the vowel /a/ during the course of development for the three groups: children (green), females (red), and males (blue).



**Figure 18.** Average  $f_0$  (across vowels) of the studies listed in Appendix A as a function of age. The three groups—children (green), females (red), and males (blue)—are the same studies as in the average plots of Figures 2–7; additional  $f_0$  data (magenta) are from studies with younger infants and children and include CS (Colton & Steinschneider, 1980), GRS (Gardosik, Ross, & Singh, 1980), KM (Kent & Murray, 1982), and KMe (Kuhl & Meltzoff, 1996). H = Hodge (1989); AK = Assmann and Katz, 2000; PO = Perry, Ohde, and Ashmead (2001); LPN = Lee, Potamianos, and Narayanan (1999); EH = Eguchi and Hirsh (1969); BP = Busby and Plant (1995).



is sharp between the ages 12 months and 3 years so that by the age of 3–5 years, the mean in males and females is about 250 Hz. A more gradual decrease in  $f_0$  appears between ages 6 and 11 years. Sex differences in  $f_0$  are strongly evident during adolescence. The overall  $f_0$  decline from infancy to adulthood is about one octave for females and two octaves for males. However, change in level of  $f_0$  is only one part of the developmental pattern in relation to laryngeal function. At some point in development, children learn to make optimal adjustments between laryngeal and supralaryngeal actions. Wermke, Mende, Manfredi, and Brusciaglioni (2002) concluded that infants age 15–17 weeks demonstrated an increased coupling and tuning between cry melody and resonance frequencies. This observation was interpreted to reflect intentional articulatory activity, that is, at this age, infants begin to make articulatory adjustments to effect greater coupling between source and vocal tract.

Lee et al. (1999) observed that  $f_0$  differences between male and female children were statistically significant beginning with the age of 12 years. However, Hacki and Heitmuller (1999) reported a lowering of both the habitual pitch and the entire speaking pitch range between the ages of 7 and 8 years for girls and between the ages of 8 and 9 years for boys. Hacki and Heitmuller also concluded that the beginning of the mutation occurs at the ages of 10–11 years. Mean  $f_0$  change is pronounced in male adolescents between the ages of about 12–15 years. For example, Lee et al. reported a 78% decrease in  $f_0$  for males between these ages. No significant change was observed after the age of 15 years, which indicates that the voice change is effectively complete by that age (Busby & Plant, 1995; Hollien, Green, & Massey, 1994; Kent & Vorperian, 1995). In summary, major developmental features of the larynx include the following: (a) substantial growth of laryngeal structures in puberty (Kent & Vorperian, 1995), (b) a lack of sexual dimorphism of the larynx in childhood (Eckel et al., 1999); and (c) differentiation of the layers of the lamina propria at about 12 years of age (Hirano, Kurita, & Nakashima, 1983; Yamashita, 1997).

## Vocal $f_0$ Data From Database Sources

Data on  $f_0$  are restricted to those studies that reported both  $f_0$  and formant results. Figure 18 shows the average  $f_0$  data across the four corner vowels as a function of age. These data mirror those reported in Kent (1976) in showing a relatively stable  $f_0$  during the 1st year, a relatively rapid decrease in early childhood, a more gradual decrease until puberty, and then a rapid decrease during adolescence (more so in males than females), whereby conspicuous differences in male–female  $f_0$  are evident by age 12 years (Lee et al., 1999; Perry et al.

2001). An implication of these data for estimates of formant frequency is that the error of estimation related to  $f_0$  should be relatively stable over the age range of about 3–12 years. It should be noted that there is substantial variability in the  $f_0$  values in different studies of infant vocalization. For example, Kuhl and Meltzoff (1996) reported a mean  $f_0$  of about 320 Hz for 12-, 16-, and 20-week-old infants who imitated vowel sounds produced by an adult model. This value is low compared with studies of infant cry and comfort-state vocalizations, perhaps because the infants in the imitation task imitated not only vowel quality but also characteristics of the speaker's voice.

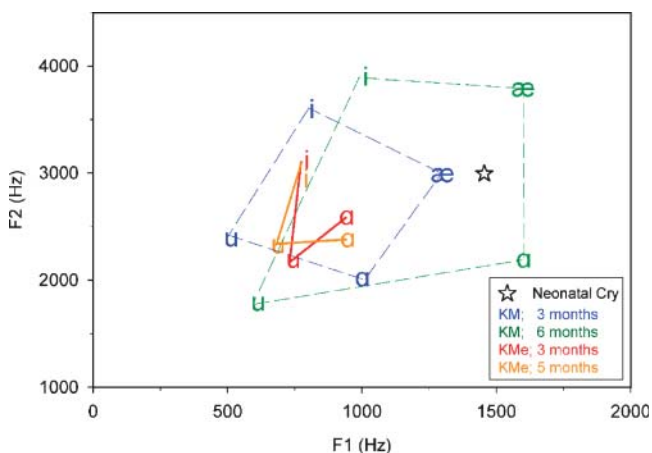
## Effects of Vocalization Type and Task

The data presented to this point pertain to vocalic segments derived from either babbling or from selected speech samples. The question arises as to how these data relate to data on other types of vocalization, such as infants' imitations of adult vowels or the vocalic elements in a newborn's cry. Establishing relationships across these different types of vocalization is a major step in understanding the developmental coherence of formant frequency data. The developmental progression seen earlier in the F1–F2 and F1–F3 patterns are the result of several factors, principally the anatomic growth of the vocal tract, the refinement of speech motor control, and the establishment of internal representations for the vocal tract configurations for the vowels of English. The interplay among these factors accounts for the results that are associated with different kinds of studies, especially when different vocalization tasks are involved.

Figure 19 shows the results of a vowel imitation study (Kuhl & Meltzoff, 1996) compared with the results of a study of spontaneous vocalizations in infancy (Kent and Murray, 1982). The imitation study analyzed the F1–F2 patterns associated with infants' imitations of the vowels /i/, /u/, and /a/ modeled by an adult speaker. The infants' F1–F2 patterns show a vowel distinctiveness in the expected directions of acoustic contrast (e.g., relatively high F2 frequency for vowel /i/), but the overall differences in the F1 and F2 frequencies are very conservative compared with the formant frequency values reported for spontaneous productions in Kent and Murray's (1982) study.

The birth cry is another vocalization type that has been studied fairly extensively. This vocalization typically marks the beginning of a lifetime of vocal behavior. In a study of 55 male and 53 female newborns, Gardosik, Ross, and Singh (1980) determined that the birth cry has an average  $f_0$  of about 460 Hz, first formant frequency of about 1550 Hz, and second formant frequency of 3100 Hz. Similarly, in a study of the cries of 35 male and 31 female newborns, Colton and Steinschneider (1980) reported a

**Figure 19.** Comparison of acoustic space of young infants from two studies—KM (Kent & Murray, 1982) and KMe (Kuhl & Meltzoff, 1996)—with different methodologies at ages 0.25 years (3 months), 0.42 years (5 months), and 0.50 years (6 months). KM used free vocalizations; KMe used imitation. The average F1–F2 data for the infant cry from Gardosik, Ross, and Singh (1980); Colton and Steinschneider (1980); and Robb and Cacace (1995) are also plotted for comparison.



mean  $f_0$  of about 510 Hz, a mean F1 frequency of about 1620 Hz, a mean F2 frequency of about 3250 Hz, and a mean F3 frequency of about 5350 Hz. Robb and Cacace (1995) compared three different methods of formant estimation (sound spectrography, linear predictive coding, and power spectrum) in the analysis of cries from 20 term infants. Estimates of F1, F2, and F3 differed somewhat with method of analysis. Means calculated from the three different methods are as follows: F1 = 1196 Hz, F2 = 2634 Hz, and F3 = 4217 Hz. The average  $f_0$  was 512 Hz. In Figure 19, the average F1 (1455 Hz) and F2 (2995 Hz) values of the infant cry from those three studies are plotted to allow comparison with the formant patterns described earlier for vowel imitations and spontaneous vowel productions. When compared with the F1–F2 and F1–F3 plots in this article, these values are relatively high and are therefore consistent with a very short vocal tract in the neonate.

## Acoustic Correlates of Velopharyngeal Anatomy

### Anatomic–Physiologic Considerations

A major allophonic variation of vowels in English is nasalization, which typically occurs when a vowel is adjacent to a nasal consonant. The capability for nasal versus nonnasal vowel production emerges in infancy, and anatomy of the velopharyngeal complex is one factor that accounts for developmental changes in nasalization.

During development, infants transition from almost exclusively nasalized vocalizations to vocalizations with an increasing degree of oral resonance. The velopharynx is open for the birth cry (Bosma, Truby, & Lind, 1965) but is closed for oral sounds by the age of 3 years (Leeper, Tissington, & Munhall, 1998; Thompson & Hixon, 1979). Although only limited data have been published for the interval between birth and 3 years, it appears that velopharyngeal closure for speech-like utterances is still developing at 6 months of age (Thom, Hoit, Hixon, & Smith, 2005), which just precedes the typical onset of canonical babbling at 7–10 months (Oller, 2000). Anatomic changes occurring around this period include a separation of the epiglottis and velum that accompanies the descent of the laryngeal framework (Sasaki, Levine, Laitman, & Crelin, 1977). This epiglottal descent continues into adolescence (Schwartz & Keller, 1997). Using MRI data, Vorperian et al. (2005, p. 342) reported data on the continuous descent of the larynx and the hyoid bone between the ages birth to 7 years, with the rate of descent being faster during the first 2 years of life.

Around the age of 3–5 years, another anatomic change may cause adjustments in velopharyngeal function. At about this time, hypertrophy of the nasopharyngeal tonsil (adenoid) is common. In an MRI study, Jaw, Sheu, Liu, and Lin (1999) reported that adenoids could be identified in only 18% of infants under the age of 3 months, 75% of infants 4 months of age, and 100% of infants older than 5 months. After rapid development in infancy, adenoids reached a plateau between 2 and 14 years of age, when they had a thickness ranging from 10.7 to 12.2 mm. After the age of 15 years, the adenoids regressed rapidly. Similar data were reported by Vogler, Li, and Pilgram (2000), who studied 189 subjects using MRI. Their data show that the adenoid pad achieved its maximum thickness (14.6 mm) during the age interval of 7–10 years. By comparison, the thickness was only about 5 mm by the age of 60 years. Vilella, Vilella, and Koch (2006) reported that adenoid sagittal thickness reached its maximum at the age of 4–5 years and progressively decreased after that age except for a slight increase at 10–11 years. Although the data from these reports are not completely congruent with respect to the age of maximum thickness of the adenoid pad, they are consistent with a lymphatic growth pattern that reaches its maximum during childhood and then follows an atrophic decline into adulthood.

In contrast to the adenoid, growth of the velopharyngeal tissues continues through adolescence. Akguner (1999) determined that growth of the hard palate ceases by the age of 15 years but that the soft palate continues to grow. Age-related anatomic changes in the velopharyngeal system may require adjustments in motor control as a child tries to maintain speech of adequate intelligibility and good quality. It has been reported that a majority of normally speaking children change their

patterns of velopharyngeal valving between prepuberty and postpuberty (Siegel-Sadewitz & Shprintzen, 1986).

## Acoustic Considerations

This section addresses two fundamental questions concerning nasality (a perceptual attribute) or nasalization (an acoustic property). The first is whether nasality or nasalization changes with development, and the second is whether speaker sex differences occur in the degree of nasality or nasalization at any age. With respect to the first question, the evidence is mixed, with some acoustic studies showing a developmental effect (Awan, 2001) but others not (Van Doorn & Purcell, 1998). When a developmental effect is observed, the degree of nasality is greater in adults than in young children. This age difference is consistent with the relatively larger lymphatic tissue in the velopharynx of children compared with adults (see preceding section and the review by Kent & Vorperian, 1995). It has been reported that nasality and duration distinguish early syllabic versus vocalic utterances, with syllabic vocalizations being longer and less nasal than vocalic ones (Bloom, 1988; Bloom, Russell, & Wassenberg, 1987). Further, Masataka and Bloom (1994) reported that adults prefer infant vocalizations that are less nasal and suggested that this preference is cross-linguistically universal. Judging from the physiologic data considered earlier, it is likely that the capability for reliable velopharyngeal closure during vocalization is developed by 7–9 months, when repetitive or canonical babbling usually begins. Sexual dimorphism is indicated by Bloom, Moore-Schoenmakers, and Masataka (1999), who reported on sex differences in the nasality of early vocalizations. Their study shows that adults rated the vocalizations of 3-month-old boys as more socially favorable (pleasant, friendly, fun, likeable, cuddly, cute) than those of girls producing similar syllabic sounds. Acoustic analyses indicated that only the feature of nasality appeared to distinguish the boys' and girls' vocalizations. They concluded that the adults' favorability ratings related to sex of the infants reflected by the less nasal acoustic quality of the boys' voices.

Whether sex-related differences in nasality extend beyond infancy is not easily answered because published studies do not present a consistent picture. Generally, acoustic studies of older children and adults have not shown sex-related differences in nasalization (Litzaw & Dalston, 1992; Mra, Sussman, & Fenwick, 1998; Prathanee, Thanaviratananich, Pongjunyakul, & Rengpatanakij, 2003; Sweeney, Sell, & O'Regan, 2004; Van Doorn & Purcell, 1998). However, in some studies, women were reported to be more nasal or have greater nasalance than men (Bloom, Zajac, & Titus, 1999; Seaver, Dalston, Leeper, &

Adams, 1991; Van Lierde, Wuyts, De Brodt, & Van Cauwenberge, 2001). Nasality is of interest for technical, biological, and cultural reasons. Technically, nasality can interfere with acoustic estimates of vowel formant frequencies because nasal resonance can be considered as a distortion added to the oral resonance pattern. If nasalization is suspected, particular care should be exercised in using linear predictive coding analyses, most of which are based on an all-pole model and neglect zeroes that arise with bifurcation of the resonator. Biologically, nasality that differs between males and females could reflect sexual dimorphism in velopharyngeal anatomy and physiology. Culturally, nasality differences between males and females could be the result of learned differences in velopharyngeal function, even if the anatomy of this system is not sexually dimorphic. To our knowledge, anatomic differences in the velopharyngeal system have not been demonstrated for children as young as 3 months. Unless the aforementioned differences in nasality between boys and girls are based on functional differences in the control of the velopharyngeal system, the most reasonable hypothesis is that undiscovered differences in velopharyngeal anatomy account for the nasality differences between infant boys and girls. It has been shown that men and women have different patterns of velopharyngeal closure (McKerns & Bzoch, 1970), but the age of appearance of this difference is not known.

---

## General Discussion

To summarize findings, acoustic data from the studies reviewed in this article indicate that vowel development is expressed as: (a) establishment of a language-appropriate acoustic representation (e.g., the F1–F2 quadrilateral or a F1–F2–F3 space), with the F1–F3 patterns having a greater developmental dispersion than the F1–F2 patterns particularly for males, and thus F1–F3 analyses may be more sensitive to changes due to age and possibly gender; (b) gradual reduction in formant frequencies with age accompanied by a decrease in F1–F2 area; (c) reduction in formant-frequency variability, possibly with an earlier stability for F1 than F2; (d) emergence of male–female differences in formant frequency by the age of 4 years, with the differences becoming more apparent by 8 years and most discrete by age 16 years; (e) nonlinear change in formant frequencies with age, with jumps in formant frequency at ages corresponding to anatomic growth spurts in all or part of the vocal tract; (f) a decline of  $f_0$  after the 1st year of life, with the decline being more rapid during early childhood (birth to 3 years) and adolescence, particularly for males whereby distinct male–female differences in  $f_0$  emerge after age 12 years—the  $f_0$  seems to be relatively stable over the age range of

about 3–12 years; (g) maturation of velopharyngeal function by about 1 year of age, which enables non-nasal vowel production; and (h) identification of speaker sex-related difference before age 12 years is mostly due to differences in the resonator but not the length of the vocal tract. The data summarized here provide a developmental perspective on one of the most frequently reported acoustic measures of speech production. These data, though limited between the ages of birth and 3 years, are a useful referent for studies of phonetic development, speaker normalization, sex differences, and other aspects of speech production.

In efforts to document vowel mastery acoustically, both the chronological age and the sex of the child should be noted. In addition, other indexes of growth—such as head circumference, neck diameter, weight, and height and percentile growth—should also be secured because height has been closely correlated to VTL (Fitch, 2000). Ideally, acoustic documentation of vowel mastery should include most of the vowels present in a particular language with special attention given to include the extreme corner vowels in the F1–F2–F3 acoustic space. The acoustic analysis also should take into account changes beyond the first two formant frequencies, should document fundamental frequency measures per vowel, and should consider formant bandwidths and assessment of nasalization. Repetition of vowel tokens secured should also be included to help assess and/or delineate variability versus developmental change. Progress in the acoustic analysis of children's speech is giving a more complete picture of factors in speech development. In particular, the data help to define the overall pattern of growth and development of the speech production system and how this pattern differs between males and females. Advancement of developmental articulatory models, such as Menard et al. (2004), that are also sex specific would also be helpful. Sexual dimorphism of the speech production system may begin in some respects in infancy and then unfolds over several years, with anatomic and physiologic differences appearing at different times in the velopharyngeal system, VTL, and the laryngeal system. Different conclusions have been published regarding the onset of sexual dimorphism of the speech production system, probably because different studies have focused on different aspects and parts of the system. It appears that a complex chronology of emerging sexual dimorphism is the most accurate picture. With the further accumulation of acoustic and anatomic data, it should be possible to construct a more accurate picture of sex differences. Another challenge is to relate formant-frequency data on vowels with imaging data on the vocal tract. As discussed in this article, the data on formant frequencies, especially for F1 and F2 alone, do not always match with conclusions derived from anatomic studies using imaging methods, such as MRI. More comprehensive

acoustic studies are needed, preferably including data on at least the first three formants (frequencies and bandwidths). It is also desirable to obtain more accurate information on the vocal tract, including three-dimensional shape and characteristics of the piriform sinuses (Baer, Gore, Gracco, & Nye, 1991; Clement et al., 2007; Dang & Honda, 1997; Story, Titze, & Hoffman, 1998).

Consideration of these data in some kind of perceptually motivated transformation is a logical next step. The current effort assembled the data in keeping with the data collection standard, which is the linear frequency scale. To be sure, various transformations would reduce to some degree the variation in formant frequencies across different age–sex groups. Selection of the ideal transform for normalization purposes is beyond the scope of this article. Several issues arise in selecting the ideal transform, including whether the procedure should be vowel-intrinsic or vowel-extrinsic (Adank, Smits, & van Hout, 2004), nature of the vowel system (Disner, 1980), and transformation algorithm (Hermansky, 1990; Hillenbrand & Houde, 2003; Miller, Engbretson, & Vemula, 1980; Syrdal & Gopal, 1986; Zahorian & Jagharghi, 1991).

The data summarized here are one step in the acoustic description of speech development, and they can be considered as a framework for the eventual acoustic description of consonants and prosodic features and for the specification of the acoustic correlates of speech disorders in children. Generalization to other languages should be done cautiously, given evidence that vowels that are considered to be phonetically equivalent in two different languages may have distinctive formant-frequency patterns (Kent & Read, 2002). It should be reiterated that dialectal influences in these data cannot be identified with certainty, but such influences likely exist. Above all, the point to be made is that research over 5 decades since the publication of the seminal article by Peterson and Barney (1952) has given a sharper, more detailed picture of the ways in which age and sex determine the formant-frequency patterns for the vowels of English. This 50-year retrospective is accompanied by the recent availability of high-quality images of the vocal tract through the methods of MRI and CT.

## Acknowledgments

This work was supported, in part, by Research Grants R03 DC4362 (Anatomic Development of the Vocal Tract: MRI Procedures), R01 DC6282 (MRI and CT Studies of the Developing Vocal Tract), and R01 DC00319 (Intelligibility Studies of Dysarthria) from the National Institute on Deafness and Other Communication Disorders. This work was also supported by National Institute of Child Health and Human Development Core Grant P-30 HD03352, awarded to the

Waisman Center. We thank Ralph Ohde for the age and vowel specific data from the Perry, Ohde, and Ashmead (2001) publication. We also thank Mary Lindstrom for preparation of Figures 2–7 and Hetal Pathak, Mike Schimek, Andrea Kettler, Allison Carolan, and Reid Durtschi for assistance with the preparation of the remaining figures. We give special thanks to Hetal Pathak and Andrea Kettler for their assistance with preparation of the summary acoustic spreadsheet from the various articles.

## References

- Adank, P., Smits, R., & van Hout, R.** (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America*, *116*, 3099–3107.
- Akguner, M.** (1999). Velopharyngeal anthropometric analysis with MRI in normal subjects. *Annals of Plastic Surgery*, *43*, 142–147.
- Arens, R., McDonough, J. M., Corbin, A. M., Hernandez, M. E., Maislin, G., Schwab, R. J., & Pack, A. I.** (2002). Linear dimensions of the upper airway structure during development: Assessment by magnetic resonance imaging. *American Journal of Respiratory and Critical Care Medicine*, *165*, 117–122.
- Assmann, P. F., Ballard, W., Bornstein, L., & Paschall, D. D.** (1994). Trackdraw: A graphical interface for controlling the parameters of a speech synthesizer. *Behavior Research Methods, Instruments, & Computers*, *26*, 431–436.
- Assmann, P. F., & Katz, W. F.** (2000). Time-varying spectral change in the vowels of children and adults. *The Journal of the Acoustical Society of America*, *108*, 1856–1866.
- Awan, S. N.** (2001). Age and gender effects on measures of RMS nasalance. *Clinical Linguistics and Phonetics*, *15*, 117–122.
- Baer, T., Gore, J. C., Gracco, L. C., & Nye, P. W.** (1991). Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels. *The Journal of the Acoustical Society of America*, *90*, 799–828.
- Bennet, S.** (1981). Vowel formant frequency characteristics of preadolescent males and females. *The Journal of the Acoustical Society of America*, *69*, 231–238.
- Bloom, K.** (1988). Quality of adult vocalizations affects the quality of infant vocalizations. *Journal of Child Language*, *15*, 469–480.
- Bloom, K., Moore-Schoenmakers, K., & Masataka, N.** (1999). Nasality of infant vocalizations determines gender bias in adult favorability. *Journal of Nonverbal Behavior*, *23*, 219–236.
- Bloom, K., Russell, A., & Wassenberg, K.** (1987). Turn taking affects the quality of infant vocalizations. *Journal of Child Language*, *14*, 211–227.
- Bloom, K., Zajac, D. J., & Titus, J.** (1999). The influence of nasality of voice on sex-stereotyped perceptions. *Journal of Nonverbal Behavior*, *23*, 271–281.
- Bosma, J. F., Truby, H. M., & Lind, J.** (1965). Cry motions of the newborn infant [Monograph]. *Acta Paediatrica Scandinavica*, *163*, 63–91.
- Buhr, R. D.** (1980). The emergence of vowels in an infant. *Journal of Speech and Hearing Research*, *23*, 73–94.
- Busby, P. A., & Plant, G. L.** (1995). Formant frequency values of vowels produced by preadolescent boys and girls. *The Journal of the Acoustical Society of America*, *97*, 2603–2606.
- Casal, C., Dominguez, C., Fernandez, A., Sarget, R., Martínez-Celdran, E., Sentis-Vilalta, J., & Gay-Escoda, C.** (2002). Spectrographic measures of the speech of young children with cleft lip and cleft palate. *Folia Phoniatrica et Logopaedica*, *54*, 247–257.
- Childers, D. G., & Wu, K.** (1991). Gender recognition from speech: Part II. Fine analysis. *The Journal of the Acoustical Society of America*, *90*, 1841–1856.
- Clement, P., Hans, S., Hartl, D. M., Maeda, S., Vaissiere, J., & Brasnu, D.** (2007). Vocal tract area function for vowels using three-dimensional magnetic resonance imaging: A preliminary study. *Journal of Voice*, *21*, 522–530.
- Clopper, C. G., Pisoni, D. B., & de Jong, K.** (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *The Journal of the Acoustical Society of America*, *118*, 1661–1676.
- Colton, R. H., & Steinschneider, A.** (1980). Acoustic relationships of infant cries to the sudden infant death syndrome. In T. Murry & J. Murry (Eds.), *Infant communication: Cry and early speech* (pp. 183–208). Houston, TX: College-Hill Press.
- Crelin, E. S.** (1973). *Functional anatomy of the newborn*. New Haven, CT: Yale University Press.
- Dang, J., & Honda, K.** (1997). Acoustic characteristics of the piriform fossa in models and humans. *The Journal of the Acoustical Society of America*, *101*, 456–465.
- De Wet, F., Weber, K., Boves, L., Cranen, B., Bengio, S., & Burlard, H.** (2004). Evaluation of formant-like features in an automatic vowel classification task. *The Journal of the Acoustical Society of America*, *116*, 1781–1792.
- Disner, S. F.** (1980). Evaluation of vowel normalization procedures. *The Journal of the Acoustical Society of America*, *67*, 253–261.
- Donegan, P.** (2002). Normal vowel development. In M. J. Ball & F. Gibbon (Eds.), *Vowel disorders* (pp. 1–35). Boston: Butterworth/Heinemann.
- Eckel, H. E., Koebke, J., Sittel, C., Sprinzl, G. M., Potoschnig, C., & Stennert, E.** (1999). Morphology of the human larynx during the first five years of life studied on whole organ serial sections. *Annals of Otolaryngology, Rhinology, & Laryngology*, *108*, 232–238.
- Eckel, H. E., Sprinzl, G. M., Sittel, C., Koebke, J., Damm, M., & Stennert, E.** (2000). Zur anatomie von glottis und subglottis beim kindlichen kehlkopf [Anatomy of the vocal folds and subglottic airway in children]. *HNO*, *48*, 501–507.
- Eguchi, S., & Hirsh, I. J.** (1969). Development of speech sounds in children. *Acta Otolaryngologica*, *257*, 5–51.
- Endres, W., Bambach, W., & Flosser, G.** (1971). Voice spectrograms as a function of age, voice disguise, and voice imitation. *The Journal of the Acoustical Society of America*, *49*, 1842–1848.
- Epps, J., Dowd, A., Smith, J., & Wolfe, J.** (1997). Real time measurements of the vocal tract resonances during speech.

- In Eurospeech97 (pp. 721–724). Rhodes, Greece: European Speech Communication Association.
- Fant, G.** (1975). A note on vocal tract size factors and non-uniform F-pattern scalings. *Speech Transmission Laboratory Quarterly Progress & Status Reports*, 4, 22–30.
- Ferguson, C. A., & Farwell, C. B.** (1975). Words and sounds in early language acquisition. *Language*, 51, 419–439.
- Fitch, W. T.** (2000). The evolution of speech: A comparative review. *Trends in Cognitive Sciences*, 4, 258–267.
- Fitch, W. T., & Giedd, J.** (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, 106, 1511–1522.
- Fort, A., & Manfredi, C.** (1998). Acoustic analysis of newborn infant cry signals. *Medical Engineering & Physics*, 20, 432–442.
- Gardosik, T. A., Ross, P. J., & Singh, S.** (1980). In T. Murry & J. Murry (Eds.), *Infant communication: Cry and early speech* (pp. 106–123). Houston, TX: College-Hill Press.
- Gilbert, H. R., Robb, M. P., & Chen, Y.** (1997). Formant frequency development—15 to 36 months. *Journal of Voice*, 11, 260–266.
- Green, J. R., Moore, C. A., & Reilly, K. U.** (2002). The sequential development of jaw and lip control in speech. *Journal of Speech, Language, and Hearing Research*, 45, 66–79.
- Hacki, T., & Heitmuller, S.** (1999). Development of the child's voice: Premutation, mutation. *International Journal of Pediatric Otorhinolaryngology*, 49(Suppl. 1), S141–S144.
- Hagiwara, R. E.** (1997). Dialect variation and formant frequency: The American English vowels revisited. *The Journal of the Acoustical Society of America*, 102, 655–658.
- Hammond, T. H., Gray, S. D., & Butler, J.** (2000). Age- and gender-related collagen distribution in human vocal folds. *Annals of Otolaryngology, Rhinology, & Laryngology*, 109, 913–920.
- Hammond, T. H., Gray, S. D., Butler, J., Zhou, R., & Hammond, E.** (1998). Age- and gender-related elastin distribution changes in human vocal folds. *Otolaryngology-Head & Neck Surgery*, 119, 314–322.
- Hermansky, H.** (1990). Perceptual linear predictive (PLP) analysis of speech. *The Journal of the Acoustical Society of America*, 87, 1738–1752.
- Higgins, C. M., & Hodge, M. M.** (2001). F2/F1 vowel quadrilateral area in young children with and without dysarthria. *Canadian Acoustics*, 29, 66–68.
- Hillenbrand, J. M., Getty, L. A., Clark, M. J., & Wheeler, K.** (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, 97, 3099–4111.
- Hillenbrand, J. M., & Houde, R. A.** (2003). A narrow band pattern-matching model of vowel perception. *The Journal of the Acoustical Society of America*, 113, 1044–1055.
- Hirano, M., Kurita, S., & Nakashima, T.** (1983). Growth, development and aging of human vocal folds. In D. M. Bless & J. H. Abbs (Eds.), *Vocal fold physiology* (pp. 22–43). San Diego, CA: College-Hill Press.
- Hodge, M.** (1989). *A comparison of spectral-temporal measures across speaker age: Implications for an acoustic characterization of speech maturation*. Unpublished doctoral dissertation, University of Wisconsin—Madison.
- Hollien, H., Green, R., & Massey, K.** (1994). Longitudinal research on adolescent voice change in males. *The Journal of the Acoustical Society of America*, 96, 2646–2654.
- Irwin, J. V., & Wong, S. P.** (1983). *Phonological development in children 18 to 72 months*. Carbondale, IL: Southern Illinois University Press.
- Israel, H.** (1968). Continuing growth in the human cranial skeleton. *Archives of Oral Biology*, 13, 133–137.
- Israel, H.** (1973). Age factor and the pattern of change in craniofacial structures. *American Journal of Physical Anthropology*, 39, 111–128.
- Ives, D. T., Smith, D. R., & Patterson, R. D.** (2005). Discrimination of speaker size from syllable phrases. *The Journal of the Acoustical Society of America*, 118, 3816–3822.
- Jaw, T. S., Sheu, R. S., Liu, G. C., & Lin, W. C.** (1999). Development of adenoids: A study by measurement with MR images. *Kaohsiung Journal of Medical Sciences*, 15, 12–18.
- Kazarian, A. G., Sarkissian, L. S., & Isaakian, D. G.** (1978). Length of the human vocal cords by age. *Zhurnal Eksperimentalnoi I Klinicheskoi Meditsiny*, 18, 105–109.
- Kent, R. D.** (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *Journal of Speech and Hearing Research*, 19, 421–447.
- Kent, R. D.** (1994). *Reference manual for communicative sciences and disorders: Speech and language*. Austin, TX: Pro-Ed.
- Kent, R. D., & Murray, A. D.** (1982). Acoustic features of infant vocalic utterances. *The Journal of the Acoustical Society of America*, 72, 353–365.
- Kent, R. D., Netsell, R., Osberger, M. J., & Hustedde, C. G.** (1987). Phonetic development in twins who differ in auditory function. *Journal of Speech and Hearing Disorders*, 52, 64–75.
- Kent, R. D., & Read, C.** (2002). *The acoustic analysis of speech* (2nd ed.). Albany, NY: Singular/Thomson Learning.
- Kent, R. D., & Vorperian, H. K.** (1995). Anatomic development of the craniofacial-oral-laryngeal systems: A review. *Journal of Medical Speech-Language Pathology*, 3, 145–190.
- Kuhl, P. K., & Meltzoff, A. N.** (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *The Journal of the Acoustical Society of America*, 100, 2425–2438.
- Lee, S., Potamianos, A., & Narayanan, S.** (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *The Journal of the Acoustical Society of America*, 105, 1455–1468.
- Leeper, H. A., Tissington, M. L., & Munhall, K. G.** (1998). Temporal aspects of velopharyngeal function in children. *Cleft Palate-Craniofacial Journal*, 35, 215–221.
- Lieberman, D. E., McCarthy, R. C., Hiiemae, K. M., & Palmer, J. B.** (2001). Ontogeny of postnatal hyoid and larynx descent in humans. *Archives of Oral Biology*, 46, 117–128.
- Linville, S. E., & Rens, J.** (2001). Vocal tract resonance analysis of aging voice using long-term average spectra. *Journal of Voice*, 15, 323–330.
- Litzaw, L. L., & Dalston, R. M.** (1992). The effect of gender upon nasalance scores among normal adult speakers. *Journal of Communication Disorders*, 25, 55–64.

- Liu, H. M., Tsao, F. M., & Kuhl, P. K.** (2005). The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *The Journal of the Acoustical Society of America*, *117*, 3879–3889.
- Maeda, S.** (1979). An articulatory model of the tongue based on a statistical analysis. *The Journal of the Acoustical Society of America*, *65*, S22.
- Maeda, S.** (1990). Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In W. L. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 131–149). Dordrecht, the Netherlands: Kluwer Academic.
- Martland, P., Whiteside, S. P., Beet, S. W., & Baghai-Ravary, L.** (1996, October). Estimating child and adolescent formant frequency values from adult data. *Proceedings of the Applied Science and Engineering Laboratories Conference ICSLP'96*, 622–625.
- Masataka, N., & Bloom, K.** (1994). Acoustic properties that determine adults' preferences to 3-month-old infant vocalizations. *Infant Behavior and Development*, *17*, 461–464.
- Mathworks.** (1993). MATLAB: Higher performance numeric computation and visualization software [Computer program]. Natick, MA: Author.
- McKerns, D., & Bzoch, K.** (1970). Variations in velopharyngeal valving: The factor of sex. *Cleft Palate Journal*, *7*, 652–662.
- Meddis, R., & Hewitt, M.** (1991). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. *The Journal of the Acoustical Society of America*, *89*, 2866–2882.
- Menard, L., Schwartz, J.-L., & Boe, L.-J.** (2004). Role of vocal tract morphology in speech development: Perceptual targets and sensorimotor maps for synthesized French vowels from birth to adulthood. *Journal of Speech, Language, and Hearing Research*, *47*, 1059–1080.
- Miller, J. D., Engebretson, A. M., & Vemula, N. R.** (1980). Vowel normalization: Differences between vowels spoken by children, women, and men. *The Journal of the Acoustical Society of America*, *68*, S33.
- Molis, M. R.** (2005). Evaluating models of vowel perception. *The Journal of the Acoustical Society of America*, *118*, 1062–1071.
- Moura, C. P., Cunha, L. M., Vilarinho, H., Cunha, M. J., Freitas, D., Palha, M., et al.** (in press). Voice parameters in children with Down syndrome. *Journal of Voice*.
- Mra, Z., Sussman, J. E., & Fenwick, J.** (1998). HONC measures in 4- to 6-year-old children. *Cleft Palate-Craniofacial Journal*, *35*, 408–414.
- Munson, B., & Solomon, P. N.** (2004). The effect of phonological neighborhood density on vowel articulation. *Journal of Speech, Language, and Hearing Research*, *47*, 1048–1058.
- Nijland, L., Maassen, B., Van der Meulen, S., Gabreels, F., Kraaiaat, F. W., & Schreuder, R.** (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics and Phonetics*, *16*, 461–483.
- Nishimura, T.** (2003). Comparative morphology of the hyolaryngeal complex in anthropoids: Two steps in the evolution of the descent of the larynx. *Primates*, *44*, 41–49.
- Nishimura, T., Mikami, A., Suzuki, J., & Matsuzawa, T.** (2006). Descent of the hyoid in chimpanzees: Evolution of face flattening and speech. *Journal of Human Evolution*, *51*, 244–254.
- Nittrouer, S.** (1993). The emergence of mature gestural patterns is not uniform: Evidence from an acoustic study. *Journal of Speech and Hearing Research*, *36*, 959–972.
- Oller, D. K.** (2000). *The emergence of the speech capacity*. Mahwah, NJ: Erlbaum.
- Pentz, A., & Gilbert, H.** (1983). *Comparison of formants in preadolescent children's vowel productions*. Poster session presented at the 1983 Annual Convention of the American Speech-Language-Hearing Association, Cincinnati, OH.
- Perry, T. L., Ohde, R. N., & Ashmead, D. H.** (2001). The acoustic bases for gender identification from children's voices. *The Journal of the Acoustical Society of America*, *109*, 2988–2998.
- Peterson, G. E., & Barney, H. L.** (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*, 585–594.
- Prathanee, B., Thanaviratnanich, S., Pongjunyakul, A., & Rengpatanakij, K.** (2003). Nasalance scores for speech in normal Thai children. *Scandinavian Journal of Plastic & Reconstructive Surgery & Hand Surgery*, *37*, 351–355.
- Rastatter, M. P., McGuire, R. A., Kalinowski, J., & Stuart, A.** (1997). Formant frequency characteristics of elderly speakers in contextual speech. *Folia Phoniatrica et Logopaedica*, *49*, 1–8.
- Robb, M. P., & Cacace, A. T.** (1995). Estimation of formant frequencies in infant cry. *International Journal of Pediatric Otorhinolaryngology*, *32*, 57–67.
- Robb, M. P., Chen, Y., & Gilbert, H. R.** (1997). Developmental aspects of formant frequency and bandwidth in infants and toddlers. *Folia Phoniatrica et Logopaedica*, *49*, 88–95.
- Rvachew, S., Slawinski, E. B., Williams, M., & Green, C. L.** (1996). Formant frequencies of vowels produced by infants with and without early onset otitis media. *Canadian Acoustics/Acoustique Canadienne*, *24*, 19–28.
- Sasaki, C. T., Levine, P. A., Laitman, J. T., & Crelin, E. S., Jr.** (1977). Postnatal descent of the epiglottis in man: A preliminary report. *Archives of Otolaryngology*, *103*, 169–171.
- Sato, K., Hirano, M., & Nakashima, T.** (2001). Fine structure of the human newborn and infant vocal fold mucosae. *Annals of Otolaryngology, Rhinology, & Laryngology*, *110*, 417–424.
- Schenk, B. S., Baumgartner, W. D., & Hamzavi, J. S.** (2003). Changes in vowel quality after cochlear implantation. *ORL Journal of Otorhinolaryngology Related Specialties*, *65*, 184–188.
- Schwartz, D. S., & Keller, M. S.** (1997). Maturational descent of the epiglottis. *Archives of Otolaryngology-Head and Neck Surgery*, *123*, 627–628.
- Scukanec, G. P., Petrosino, L., & Squibb, K.** (1991). Formant frequency characteristics of children, young adult, and aged female speakers. *Perceptual & Motor Skills*, *73*, 203–208.
- Seaver, E. J., Dalston, R. M., Leeper, H. A., & Adams, L. E.** (1991). A study of nasometric values for normal nasal resonance. *Journal of Speech and Hearing Research*, *34*, 715–721.

- Siegel-Sadewitz, V. L., & Shprintzen, R. J.** (1986). Changes in velopharyngeal valving with age. *International Journal of Pediatric Otorhinolaryngology*, *11*, 171–182.
- Smith, A., & Goffman, L.** (2004). Interaction of motor and language factors in the development of speech. In B. Maassen, R. Kent, H. Peters, P. van Lieshout, & W. Hulstijn (Eds.), *Speech motor control in normal and disordered speech* (pp. 227–252). Oxford, England: Oxford University Press.
- Smith, D. R., Patterson, R. D., Turner, R., Kawahara, H., & Irino, T.** (2005). The processing and perception of size information in speech sounds. *The Journal of the Acoustical Society of America*, *117*, 305–318.
- Story, B. H., Titze, I. R., & Hoffman, E. A.** (1998). Vocal tract area functions for an adult female speaker based on volumetric imaging. *The Journal of the Acoustical Society of America*, *104*, 471–487.
- Sweeney, T., Sell, D., & O'Regan, M.** (2004). Nasalance scores for normal-speaking Irish children. *Cleft Palate-Craniofacial Journal*, *41*, 168–174.
- Syrdal, A. K., & Gopal, H. S.** (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America*, *79*, 1086–1100.
- Templin, M. C.** (1957). *Certain language skills in children: Their development and interrelationships*. Minneapolis, MN: University of Minnesota Press.
- Thom, S., Hoit, J., Hixon, T., & Smith, A.** (2005). Velopharyngeal function during vocalization in infants. *Cleft Palate-Craniofacial Journal*, *43*, 539–546.
- Thompson, A. E., & Hixon, T. J.** (1979). Nasal air flow during normal speech production. *Cleft Palate Journal*, *16*, 412–420.
- Trautmüller, H., & Eriksson, A.** (1997). A method for measuring formant frequencies at high fundamental frequencies. *Proceedings of EuroSpeech '97*, *1*, 470–480.
- Van Doorn, J., & Purcell, A.** (1998). Nasalance levels in the speech of normal Australian children. *Cleft Palate-Craniofacial Journal*, *35*, 287–292.
- Van Lierde, K. M., Wuyts, F. L., De Brodt, M., & Van Cauwenberge, P.** (2001). Nasometric values for normal nasal resonance in the speech of young Flemish adults. *Cleft Palate-Craniofacial Journal*, *38*, 112–118.
- Vilella, B. D., Vilella, O. D., & Koch, H. A.** (2006). Growth of the nasopharynx and adenoidal development in Brazilian subjects. *Pesquisa Odontologica Brasileira*, *20*, 70–75.
- Vogler, R. C., Li, F. J., & Pilgram, T. K.** (2000). Age-specific size of the normal adenoid pad on magnetic resonance imaging. *Clinics in Otolaryngology and Allied Sciences*, *25*, 392–395.
- Vorperian, H. K.** (2000). *Anatomic development of the vocal tract structures as visualized by MRI*. Unpublished doctoral dissertation, University of Wisconsin—Madison.
- Vorperian, H. K., Kent, R. D., Gentry, L. R., & Yandell, B. S.** (1999). MRI procedures to study the concurrent anatomic development of the vocal tract structures: Preliminary results. *International Journal of Pediatric Otorhinolaryngology*, *49*, 197–206.
- Vorperian, H. K., Kent, R. D., Lindstrom, M. J., Kalina, C. M., Gentry, L. R., & Yandell, B. S.** (2005). Development of vocal tract length during childhood: A magnetic resonance imaging study. *The Journal of the Acoustical Society of America*, *117*, 338–350.
- Walsh, B., & Smith, A.** (2002). Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes. *Journal of Speech, Language, and Hearing Research*, *45*, 1119–1133.
- Wermke, K., Mende, W., Manfredi, C., & Brusciaglioni, P.** (2002). Developmental aspects of infant's cry melody and formants. *Medical Engineering & Physics*, *24*, 501–514.
- White, P.** (1999). Formant frequency analysis of children's spoken and sung vowels using sweeping fundamental frequency production. *Journal of Voice*, *13*, 570–582.
- Whiteside, S. P.** (2001). Sex-specific fundamental and formant frequency patterns in a cross-sectional study. *The Journal of the Acoustical Society of America*, *110*, 464–478.
- Whiteside, S. P., & Hodgson, C.** (2000). Speech patterns of children and adults elicited via a picture-naming task: An acoustic study. *Speech Communication*, *32*, 267–285.
- Xue, S. A., & Hao, J. G.** (2003). Changes in the human vocal tract due to aging and the acoustic correlates of speech production: a pilot study. *Journal of Speech, Language, and Hearing Research*, *46*, 689–701.
- Xue, S. A., & Hao, J. G.** (2006). Normative standards for vocal tract dimensions by race as measured by acoustic pharyngotomy. *Journal of Voice*, *20*, 391–400.
- Yamashita, K.** (1997). Age-related development of the arrangement of connective tissue fibers in the lamina propria of the human vocal folds—Scanning electron microscope examination with digestion method. *Nippon Jibiinkoka Gakkai Kaiho*, *100*, 495–511.
- Yang, B.** (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, *24*, 245–261.
- Zahorian, S. A., & Jagharghi, A. J.** (1991). Speaker normalization of static and dynamic vowel spectral features. *The Journal of the Acoustical Society of America*, *90*, 67–75.
- Zahorian, S. A., & Jagharghi, A. J.** (1993). Spectral-shape features versus formants as acoustic correlates for vowels. *The Journal of the Acoustical Society of America*, *94*, 1966–1982.

Received September 12, 2006

Accepted April 4, 2007

DOI: 10.1044/1092-4388(2007)104

Contact author: Houri K. Vorperian, 481 Waisman Center, 1500 Highland Avenue, Madison, WI 53705.

E-mail: vorperian@waisman.wisc.edu.

**Appendix A** (p. 1 of 9). Summary table of the 21 reviewed studies on vowel development listed chronologically.

Abbr.	Study	Subject detail	Age	Vowels in plots	Methods of obtaining the vowel	Additional details	F1–F2 plot C–M–F	F1–F3 plot C–M–F
PB*	Peterson and Barney (1952)	<i>n</i> = 76 Children: <i>n</i> 1 = 15 Adults: <i>n</i> 2 = 33 M, <i>n</i> 3 = 28 F	Children 9 years plus adults	/i/, /u/, /ae/, /a/	Lists with 10 monosyllabic words: heed, hid, head, had, hod, hawed, hood, who'd, hud, heard  General American English	Two random word lists per speaker producing 1,520 recorded words. Analysis via sound spectrograph.  Formant frequencies estimated from weighted average of the frequencies of the principal components in the formant.	C at 9 years  M and F at adults	C at 9 years  M and F at adults
EH*	Eguchi and Hirsh (1969)	<i>n</i> = 84 <i>n</i> per age/sex group = 5/6	Children 3–13 years plus adults	/i/, /u/, /ae/, /a/	Two sentences: He has a blue pen. I am tall.  Vowels in American English	Each sentence produced/read on five different occasions. Repetition after a native speaker for children <7 years. Used wide-band and narrow-band spectrograms. F1 and F2 estimated from spectrum envelopes drawn on expanded narrow-band sections (0–4000 Hz).	C at 3, 4, 5, 6, 7, 8, 9, 10, and 11 years  M and F at (a) 11, 12, and 13 years and (b) adults	No F3 values reported.
B	Buhr (1980)	<i>n</i> = 1 M	Children 16–64 weeks	/i/, /u/, /ae/, /a/	Select sounds classified by phonetician as a particular vowel sound of English	Biweekly recordings; 944 vowels analyzed. Spectrograms measured by determining the first three formants for each sound. Formant values for the different vowels can be extracted from the plots.	Not plotted	Not plotted

**Appendix A** (p. 2 of 9). Summary table of the 21 reviewed studies on vowel development listed chronologically.

Abbr.	Study	Subject detail	Age	Vowels in plots	Methods of obtaining the vowel	Additional details	F1–F2 plot C–M–F	F1–F3 plot C–M–F
KM*	Kent and Murray (1982)	n = 21 n1 = 7 (3F, 4M) n2 = 7 (3F, 4M) n3 = 7 (4F, 3M)	Children n1 = 3 months n2 = 6 months n3 = 9 months	Extreme F1–F2 values of the four corners were used to define acoustic space.	Infant vocalizations (comfort state) coded according to vocalic, phonation, and noise segment (see KM's Coding Table 1, p. 355)	Vocalic utterances: f <sub>0</sub> determined from narrow-band (45-Hz) spectrograms, and formant frequencies determined from wide-band (450-Hz) spectrograms. F1–F2 formant values per age group were extracted from KM's Figures 8–10, p. 357.	C at 9 months Note: Ages 3 and 6 months not plotted.	Average F3 = 5K. No mean F3 values for the different corner vowels reported.
PG*	Pentz and Gilbert (1983)	n = not available	Children 7, 8, and 9 years	/i/, /u/, /ae/, /a/	Not available	Not available. Poster presentation at the 1983 annual convention of the American Speech-Language-Hearing Association. Values reported in Kent (1994), p. 73.	C at 7, 8, and 9 years	C at 7, 8, and 9 years
H*	Hodge (1989)	n = 115 n1 = 15 (12 M, 3 F) n2 = 20 (19 M, 1F) n3 = 20 (20 M)	Children n1 = infants (M = 8.5 months) n2 = 1 year (M = 18 months) n3 = 3, 5, and 9 years plus adults	<9 months = /i/, /u/, /ae/ 1 year = adult /i/, /u/, /ae/, /a/	Isolated productions of the point and central vowels preceded by aspirate /h/; reduplicated triple stop consonant vowel (CV) chain; words "baby" and "bye bye"; sustained /s/ sounds; and words "two" and "tea"; and two word combinations "a stee" and "a stew"  Vowels in Canadian English	Used spectrographic displays with filter bandwidth and dynamic range settings that provided the clearest formant pattern. F1, F2, and F3 values were estimated from the vocalic segments by tracing the midpoint of the strong energy region corresponding to each of the formants.	C at 9 months and 1, 3, 5, and 9 years M at (a) 3, 5, and 9 years and (b) adults	C at 9 months and 1, 3, 5, and 9 years M at (a) 3, 5, and 9 years and (b) adults

**Appendix A** (p. 3 of 9). Summary table of the 21 reviewed studies on vowel development listed chronologically.

Abbr.	Study	Subject detail	Age	Vowels in plots	Methods of obtaining the vowel	Additional details	F1–F2 plot C–M–F	F1–F3 plot C–M–F
<b>CW*</b>	Childers and Wu (1991)	<i>n</i> = 52 (27 M, 25 F)	Adults 20–80 years	/i/, /u/, /ae/, /a/	Sustained each vowel as it would be pronounced in the following words: beet, bit, bet, bat, Bob, bought, book, boot, but, Bert American English vowels	Formant information extracted by closed phase weighted recursive least square method with a variable forgetting factor.	M and F at adults	M and F at adults
<b>ZJ*</b>	Zahorian and Jagharghi (1993)	<i>n</i> = 30 Children: <i>n</i> 1 = 10 (5 M, 5 F) Adults: <i>n</i> 2 = 20 (10 M, 10 F)	Children 11 years plus adults	/i/, /u/, /ae/, /a/	A total of 99 CVC syllables produced in isolation. CVC syllable list contained nine instances of each of the following 11 vowels: /iy/, /ih/, /eh/, /ae/, /ah/, /aa/, /ao/, /ow/, /uh/, /uw/, and /er/. CVC list is in Zahorian and Jagharghi's (1991) Appendix A. Vowels in American English—various dialects	A total of 2,922 vowels/ formant frequency analyzed; 50-ms Hamming window and 10th order linear-programming models computed. Used a dynamic programming approach to select actual formant values with formant seed values/data from Peterson and Barney (1952). Formant tracking verified visually and with global spectral shape.	C at 11 years M and F at adults	C at 11 years M and F at adults
<b>BP*</b>	Busby and Plant (1995)	<i>n</i> = 40 <i>n</i> = 10 per age group (5 M, 5 F)	Children 5, 7, 9, and 11 years	/i/, /u/, /ae/, /a/	Words: sheep, ship, bed, cat, cart, cut, four, dog, shoe, book, bird Carrier phrase: "I can see a ..." Test words presented orthographically and pictorially. Vowels in Australian English	A total of 440 values per formant frequency. <i>f</i> <sub>0</sub> estimated from spectrographic traces of the first two harmonics. F1–F3 values determined from the steady state portion of vowel using spectrogram and average power spectrum displays. Formant values extracted from BP's Figures 2–3, p. 2605.	C at 5, 7, 9, and 11 years M and F at 5, 7, 9, and 11 years	No mean F3 values per vowel. Mean F3 across all vowels per age group in BP's Figure 1, p. 2604.

**Appendix A** (p. 4 of 9). Summary table of the 21 reviewed studies on vowel development listed chronologically.

Abbr.	Study	Subject detail	Age	Vowels in plots	Methods of obtaining the vowel	Additional details	F1–F2 plot C–M–F	F1–F3 plot C–M–F
HW*	Hagiwara (1997)	<i>n</i> = 15 (6 M, 9 F)	Adults 18–26 years	/i/, /u/, /ae/, /a/	Database of 33 words with 11 vowels: beat, bit, bate, bet, bat, boot, put, boat, bought, but, Bert Each word presented in the frame “Cite ... twice.” Vowels in American English	A 30-ms window digitized at 10 KHz. Formants determined from wide-band spectrogram and narrow-band fast Fourier transform spectra and linear predictive coding (LPC) analysis.	M and F at adults	M and F at adults
HG*	Hillenbrand, Getty, Clark, and Wheeler (1995)	<i>n</i> = 139 Children: <i>n</i> 1 = 46 (27 M, 19 F) Adults: <i>n</i> 2 = 93 (45 M, 48 F)	Children 10–12 years plus adults	/i/, /u/, /ae/, /a/	Read randomized lists with the following words: heed, hid, hayed, had, hod, hawed, hoed, hood, who’d, hud, heard, hoyed, hide, hewed, how’d Vowels in American English—upper midwest dialect	Digitized at 16 kHz with 12-bit resolution. F1–F4 measured from LPC spectra over 16-ms Hamming windows. Measures made while viewing a spectral peak display and a spectrogram.	C at 11 years M and F at adults	C at 11 years M and F at adults
Y*	Yang (1996)	<i>n</i> = 20 (10 M, 10 F)  Note: The <i>n</i> reflects only the number of American English speakers/participants.	Adults 18–27 years	/i/, /u/, /ae/, /a/	Read list with the following 13 vowels/words: had, hard, hawed, hayed, head, heed, herd, hid, hod, hoed, who’d, Hudd, and hood Each word was present/produced five times. American English vowels. Korean vowels are not included in this summary.	F1–F3 are average values of three repetitions/vowels per speaker. Input samples were low-pass filtered at 4 kHz and digitized at a 10-kHz sampling rate. Spectrograms used 256-point discrete Fourier transform analysis with a 6.4-ms Hamming window once every millisecond. Formant measure taken one third into the vowel duration (on/offset).	M and F at adults	M and F at adults

**Appendix A** (p. 5 of 9). Summary table of the 21 reviewed studies on vowel development listed chronologically.

Abbr.	Study	Subject detail	Age	Vowels in plots	Methods of obtaining the vowel	Additional details	F1–F2 plot C–M–F	F1–F3 plot C–M–F
KMe	Kuhl and Meltzoff (1996)	$n = 72$ $n = 24$ per age group	Children 3, 4, and 5 months	/i/-, /u/-, and /a/-like vowels	Infants listened to eight different productions representative of /i/ as in “heap,” /u/ as in “hoop,” and /a/ as in “hop.” Infants produced vocalizations resembling the vowel they heard. Vowel-like vocalizations analyzed following perceptual coding as /i/-, /u/-, and /a/-like.	Formant values were based on agreement of three analyses: Narrow-band spectrogram (114 Hz), fast Fourier transform (256 points, preemphasis, Blackman window weighting), and LPC frequency response (10-ms frame length, filter order = 12).	Not plotted	Not plotted
GRC	Gilbert, Robb, and Chen (1997)	$n = 4$ (4 M)	Children 15–36 months Sampling sessions at 15, 18, 21, 24, and 36 months	Vowels not specified	Spontaneous vocalizations. Phonetic classifications made as belonging to tongue height (high, middle, low) and tongue advancement features (front, central, back).	A total of 1,334 vowels analyzed. Analysis used a wide-band (450-Hz) spectrogram. Center frequencies of the first two steady state formants were judged to be F1 and F2.	Not plotted	Not plotted

**Appendix A** (p. 6 of 9). Summary table of the 21 reviewed studies on vowel development listed chronologically.

Abbr.	Study	Subject detail	Age	Vowels in plots	Methods of obtaining the vowel	Additional details	F1–F2 plot C–M–F	F1–F3 plot C–M–F
RG	Robb, Chen, and Gilbert (1997)	<i>n</i> = 20 (9 M, 11 F)	Children 4–25 months	Average formant frequency	Comfort-state noncry vocalizations were collected from each child. A total of 1,743 vowel-like sounds. An average of 88 vowel-like sounds per child. Every distinguishable vocalization was phonetically transcribed.	Digitized at 10 kHz with 16-bit quantization. Using an Amplitude × Time waveform, the on/offset per vowel was determined and transformed to a narrow-band spectrogram (24 Hz). F1 and F2 were determined from LPC spectra (six coefficients across a 20-ms Hamming window).	Not plotted	Not plotted
LPN*	Lee, Potamianos, and Narayanan (1999)	<i>n</i> = 492 Children: <i>n</i> <sub>1</sub> = 436 (229 M, 207 F) Adults: <i>n</i> <sub>2</sub> = 56	Children 5–18 years Adults 25–50 years	/i/, /u/, /ae/, /a/	Fifteen target words and five sentences presented on a computer monitor. Utterances produced twice in random. Words were as follows: bead, bit, bet, bat, pot, ball, but, put, boot, bird. Words were produced in a carrier sentence: "I say uh—again"; 5–6-year-olds produced target words in isolation. Sentences were as follows: He has a blue pen. I am tall. She needs strawberry jam on her toast. Chuck seems thirsty after the race. Did you like the zoo this spring? Vowels in American English	A total of 3,265 pairs analyzed. Digitized at a 20-kHz sampling rate and 16-bit resolution. Utterances segmented using the AT&T hidden Markov model recognition engine. <i>f</i> <sub>0</sub> and formant values were secured using an automatic <i>F</i> <sub>0</sub> and formant-tracking program in an entropic signal processing system. Automatic versus a sample of manual estimations were compared. Waveforms were downsampled to 12 kHz and processed using a 12-ms Hamming window.	C at 5, 6, 7, 8, 9, 10, and 11 years M and F at (a) 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, and 18 years and (b) adults	C at 5, 6, 7, 8, 9, 10, and 11 years M and F at (a) 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, and 18 years and (b) adults

**Appendix A** (p. 7 of 9). Summary table of the 21 reviewed studies on vowel development listed chronologically.

Abbr.	Study	Subject detail	Age	Vowels in plots	Methods of obtaining the vowel	Additional details	F1–F2 plot C–M–F	F1–F3 plot C–M–F
WH	Whiteside and Hodgson (2000)	<i>n</i> = 29 Children: <i>n</i> <sub>1</sub> = 20 (10 M, 10 F) Adults: <i>n</i> <sub>2</sub> = 9 (4 M, 5 F)	Children 6 and 9 years (3 M, 3 F) Age 10 years (4 M, 4 F) Adults mean age 37.4 years	/a/	Picture naming task. Target phrases: The red/blue/green bar/jar/car. Four distracters: The red boat, The green balloon. Vowels in United Kingdom English	Nine phrase final vowels per participant analyzed. Digitized at a 10-kHz sampling rate. F <sub>0</sub> calculated using the autocorrelation method (20-ms frame length). F1–F3 calculated at the midpoint of final vowel using automatic LPC analysis.	Not plotted	Not plotted
AK*	Assmann and Katz (2000)	<i>n</i> = 50 Children: <i>n</i> <sub>1</sub> = 30 (10 per age group) Adults: <i>n</i> = 20 (10 M, 10 F)	Children 3, 5, and 7 years plus adults	/i/, /u/, /ae/, /a/	Words: heed, hid, hayed, head, had, hud, hod, hawed, herd, hoed, hood, who'd Six random repetitions per word/vowel Vowels in American English	A total of 180 vowel tokens/formant frequency analyzed. Digitized at 48 kHz and a 16-bit resolution. Waveforms were low-pass filtered and resampled at 12 kHz. f <sub>0</sub> estimates made using Meddis and Hewitt's (1991) pitch model. Formant center frequencies were tracked using a custom MATLAB (Mathworks, 1993; Assmann, Ballard, Bornstein, & Paschall, 1994) program.	C at 3, 5, and 7 years M and F at adults	C at 3, 5, and 7 years M and F at adults

**Appendix A** (p. 8 of 9). Summary table of the 21 reviewed studies on vowel development listed chronologically.

Abbr.	Study	Subject detail	Age	Vowels in plots	Methods of obtaining the vowel	Additional details	F1–F2 plot C–M–F	F1–F3 plot C–M–F
POA*	Perry, Ohde, and Ashmead (2001)	<i>n</i> = 80 <i>n</i> = 20 per age group (10 M, 10 F)	Children 4, 8, 12, and 16 years	/i/, /u/, /ae/, /a/	Seven vowels in the neutral context of /hVd/—had, head, heed, hid, hod, hud, who’d—were embedded in the carrier phrase “Say /hVd/ again.” Each /hVd/ was analyzed five of the seven times produced. Vowels in American English	A total of 2,800 vowels analyzed. Digitized at a 20-kHz sampling rate. <i>f</i> <sub>0</sub> determined with computerized speech lab (CSL) pitch extraction program. F1, F2, and F3 values measured from spectrogram at the midpoint of vowel. Also, LPC analysis was used with a 10-ms triangular window (14–20 coefficients) and the cursor placed at the midpoint of F2 stability.	C at 4 and 8 years M and F at 4, 8, 12, and 16 years	C at 4 and 8 years M and F at 4, 8, 12, and 16 years
NM	Nijland et al. (2002)	<i>n</i> = 25 controls Children: <i>n</i> 1 = 19 (14 M, 5 F) Adults: <i>n</i> 2 = 6 (0 M, 6 F) Note: The <i>n</i> reflects only the number of controls for the participants with developmental apraxia of speech.	<i>n</i> 1 = 4;11 (years; months) to 6;10 Adults 6 F	/i/, /u/, /a/	Disyllabic nonsense utterances of the type [əCV], with V as one of extreme vowels /a, i, u/ of the Dutch vowel space, and C was either a fricative (the alveolar /s/ or the velar /x/) or a stop (/b/ or /d/). All syllables were produced six times in the carrier phrase “he de ... weer” (“hey the ... again”). Vowels from native speakers of Dutch	A total of 72 utterances per subject (three vowels, four consonants, six repetitions) digitized at a 25-kHz sampling rate. F2 trajectory was used to determine utterance types. CSL was used to mark amplitude peaks. Formant values were obtained using pitch synchronous LPC analyses (triangular window; 20 components autocorrelation with a preemphasis of .950).	Not plotted in summary plots. Dutch vowels formant values fall within the age specific average acoustic space for English vowels.	Not available

**Appendix A** (p. 9 of 9). Summary table of the 21 reviewed studies on vowel development listed chronologically.

Abbr.	Study	Subject detail	Age	Vowels in plots	Methods of obtaining the vowel	Additional details	F1–F2 plot C–M–F	F1–F3 plot C–M–F
CD	Casal et al. (2002)	<i>n</i> = 22 controls for the 22 children with cleft. Note: The <i>n</i> reflects only the number of control participants.	<i>n</i> for controls: <i>n</i> = 9 at 22 months, and <i>n</i> = 13 at 33 months	/i/, /u/, /a/	Spontaneous speech in Spanish elicited to secure needed speech sounds: five vowels /a, i, u, e, o/ and four consonants /p, t, k, m/ Vowels from Spanish speakers	A total of 220 vowels per formant frequency. F1 and F2 values were measured at the midpoint of the steady state of vowels. Midpoint as determined from duration of steady states. Note: Formant values from cleft lip and/or cleft palate are not considered for comparison. Only formant values from their controls are considered.	Not plotted in summary plots. Spanish acoustic space is similar to the average acoustic space for English vowels.	Not available

Note. Study abbreviations (Abbr.) in bold and marked with an asterisk in the first column indicate the 14 studies from which formant values were used in the summary average F1–F2 plots (see Figures 2, 3, and 4) and F1–F3 plots (see Figures 5, 6, and 7). The last two columns further specify the three groups (children [C], males [M], and females [F]) and the exact ages from which the formant values were included in the averages for Figures 2–7.

**Appendix B.** Summary of studies per age and per the C (children), M (males), and F (females) groups in Figures 2–7.

Age (years)	Group C–M–F	No. of studies	Studies																
0.25	*	2	KM	KMe															
0.42	*	1		KMe															
0.50	*	1	KM																
0.71	C	1			H														
0.75	C	1	KM																
1	C	1			H														
2		0																	
3	C	3			H	EH		AK											
4	C	2				EH	POA												
	F	1					POA												
	M	1					POA												
5	C	5			H	EH		AK	BP	LPN									
	F	2							BP	LPN									
	M	2							BP	LPN									
6	C	2				EH				LPN									
	F	1								LPN									
	M	1								LPN									
7	C	5				EH		AK	BP	LPN	PG								
	F	2							BP	LPN									
	M	2							BP	LPN									
8	C	4				EH	POA			LPN	PG								
	F	2					POA			LPN									
	M	2					POA			LPN									
9	C	6			H	EH			BP	LPN	PG	PB							
	F	2							BP	LPN									
	M	2							BP	LPN									
10	C	2				EH				LPN									
	F	1								LPN									
	M	1								LPN									
11	C	3				EH			BP	LPN		ZJ	HG						
	F	5				EH			BP	LPN									
	M	3				EH			BP	LPN									
12	F	3				EH	POA			LPN									
	M	3				EH	POA			LPN									
13	F	2				EH				LPN									
	M	2				EH				LPN									
14	F	1								LPN									
	M	1								LPN									
15	F	1								LPN									
	M	1								LPN									
16	F	2					POA			LPN									
	M	2					POA			LPN									
17		0																	
18	F	1								LPN									
	M	1								LPN									
19	F	9				EH		AK		LPN	PB	ZJ	HG	CW	HW	Y			
Adults	M	9			H	EH		AK		LPN	PB	ZJ	HG	CW	HW	Y			

Note. An asterisk indicates the additional studies/ages used in Figures 18 and 19. KM = Kent and Murray (1982); KMe = Kuhl and Meltzoff (1996); H = Hodge (1989); EH = Eguchi and Hirsh (1969); POA (V in Figures 14–17) = Perry, Ohde, and Ashmead (2001); AK = Assmann and Katz (2000); BP = Busby and Plant (1995); LPN = Lee, Potamianos, and Narayanan (1999); PG = Pentz and Gilbert (1983); PB = Peterson and Barney (1952); ZJ = Zahorian and Jagharghi (1993); HG = Hillenbrand, Getty, Clark, and Wheeler (1995); CW = Childers and Wu (1991); HW = Hagiwara (1997); Y = Yang (1996).