Laryngeal Factors in Voiceless Consonant Production in Men, Women, and 5-Year-Olds

Laura L. Koenig
Haskins Laboratories
New Haven, CT

Voicing control in stop consonants has often been measured by means of voice onset time (VOT) and discussed in terms of interarticulator timing. However, control of voicing also involves details of laryngeal setting and management of sub- and supraglottal pressure levels, and many of these factors are known to undergo developmental change. Mechanical and aerodynamic conditions at the glottis may therefore vary considerably in normal populations as functions of age and/or sex. The current study collected oral airflow, intraoral pressure, and acoustic signals from normal English-speaking adults and children producing stop consonants and /h/ embedded in a short carrier utterance. Measures were made of stop VOTs, /h/ voicing and flow characteristics, and subglottal pressure during /p/ closures. Clear age and gender effects were observed for /h/: Fully voiced /h/ was most common in men, and /h/ voicing and flow data showed the highest variability among the 5-year-olds. For individual participants, distributional measures of VOT in /p t/ were correlated with distributional measures of voicing in /h/. The data indicate that one cannot assume comparable laryngeal conditions across speaker groups. This, in turn, implies that VOT acquisition in children cannot be interpreted purely in terms of developing interarticulator timing control, but must also reflect growing mastery over voicing itself. Further, differences in laryngeal structure and aerodynamic quantities may require men and women to adopt somewhat different strategies for achieving distinctive consonantal voicing contrasts.

KEY WORDS: voicing/phonation, voice onset time, children, gender differences

Voiceless consonant production requires simultaneous control of glottal and supraglottal configurations and appropriate temporal phasing between articulatory events. Numerous studies have investigated the development of stop consonant voicing in children by means of voice onset time (VOT) measures and show that adult-like VOT patterns may not be attained until puberty. This finding has often been discussed in terms of interarticulator timing control development. Other factors besides laryngeal-supralaryngeal timing are involved in voicing control, however. These include the extent of glottal ad/abduction; the shape of the glottal channel; pressure levels below, within, and above the glottis; and vocal fold tissue characteristics including stiffness or compliance. All of these factors undergo developmental change and, therefore, vary within normal populations as functions of age and sex. This suggests that consonantal voicing behavior may also vary in systematic and predictable ways with age and sex. The research reported here explores this possibility in English-speaking men, women,

1092-4388/00/4305-1211

and children by collecting data on stop VOTs and comparing these to the voicing characteristics of the glottal fricative /h/.

VOT and Development of Voicing Control

The voicing characteristics of syllable-initial stop consonants have traditionally been measured in terms of VOT, defined as the time between plosive release and the onset of voicing. Lisker and Abramson (1964) found that contrastively voiced stops in many languages showed distinct VOT distributions and that three basic categories could be identified across languages. Voiceless aspirated stops show positive (long-lag) values of VOT; voiceless unaspirated stops have VOTs at or near zero (short voicing lags); and stops with voicing beginning during the consonantal closure have negative VOTs. In American English, stressed, syllable-initial /p t k/ have long positive values of VOT, and /b d g/ have either negative or short-lag VOTs.

Developmental studies have found that children attain adult-like VOT patterns in voiceless aspirated stops at a fairly late age. Many authors have reported age differences in VOT mean values, but the direction of the effect differs across studies: Kewley-Port and Preston (1974), Macken and Barton (1980), and Zlatin and Koenigsknecht (1976) reported short mean values in children up to 6 years of age relative to adults, whereas Barton and Macken (1980), Gilbert (1977), Menyuk and Klatt (1975), and Smith (1978) found relatively long VOT means in children 4 years and younger. What has been consistent in these studies, and in data from older children as well, is that the voiceless aspirated category displays wide ranges and high variability of VOT, possibly as late as puberty (Eguchi & Hirsh, 1969; Kent & Forner, 1980; Ohde, 1985; Ostry, Feltham, & Munhall, 1984).

Direct study of laryngeal activity in voiceless aspirated stops shows that VOT varies directly with the time interval between peak laryngeal abduction and oral release (Löfqvist, 1992). Voiceless unaspirated stops, on the other hand, may be produced with either adducted or abducted vocal folds, and abduction at various times during the closure will yield similar acoustic results provided that the folds are adducted at release (Sawashima, Abramson, Cooper, & Lisker, 1970). Kewley-Port and Preston (1974), drawing on the results of Sawashima et al., argued that voiceless aspirates develop later than voiceless inaspirates, specifically because they require precise timing between glottal abduction and oral release. In contrast, voiceless inaspirates can be produced with abduction occurring at various times relative to release and do not necessarily require a laryngeal gesture at all. The voiceless aspirated category has, accordingly, been regarded as a test of articulatory timing skill, and production variability has been interpreted as a demonstration of variable interarticulator phasing.

Although the relative timing of laryngeal and supralaryngeal movements is clearly one important factor in effecting VOT differences within speakers, additional parameters are involved in determining when voicing onsets occur across speech sounds and speakers. These include at least the following: pressure levels below, within, and above the glottis; magnitude of abduction (i.e., cross-sectional glottal area); the shape of the glottal channel; vocal fold tissue characteristics; and muscle use patterns (Ishizaka & Flanagan, 1972; Löfqvist, 1992; Stevens, 1977, 1988; Titze, 1988, 1994; van den Berg, 1958). When comparing across speaker populations, these factors cannot be entirely ignored because anatomical and aerodynamic data reveal systematic age- and sex-related variation along virtually all relevant dimensions.

Age and Sex Effects in Laryngeal and Aerodynamic Quantities

During a child's first 5 or 6 years, the laryngeal cartilages descend into the pharynx and undergo dramatic positional shifts relative to each other (Crelin, 1973; Goldstein, 1980). Changes within the vocal folds continue into adolescence and include lengthening, thickening, and development of layered tissue structure (Goldstein, 1980; Hirano, Kurita, & Nakashima, 1983; Kahane, 1982). Sex effects in laryngeal structure begin to appear at puberty, when males experience disproportionately large increases in cartilage size and vocal fold length and thickness. The added length gives men, on average, greater lateral amplitudes of vocal-fold vibration. Because vocal-fold growth is most extreme in the membranous region, men also have lower ratios of cartilage to membrane (Hirano et al., 1983), and hence, lower tissue stiffnesses overall than women or prepubescent children (cf. Titze, 1989). Additional vocal-fold mass may also yield a bulge along the glottal margin, making men less likely than women to have a posterior "chink" during normal voicing (Titze, 1989).

The respiratory system develops throughout childhood and adolescence. Lung volumes, capacities, and recoil pressures and expiratory flow rates all increase with age, whereas subglottal and intraoral pressures decrease (Bernthal & Beukelman, 1978; de Troyer, Yernault, Englert, Baran, & Paiva, 1978; Hoit, Hixon, Watson, & Morgan, 1990; Mansell, Bryan, & Levison, 1977; Netsell, Lotz, Peters, & Schulte, 1994; Polgar & Weng, 1979; Stathopoulos, 1995; Stathopoulos & Weismer, 1985a, 1985b; Trullinger & Emanuel, 1983).

In contrast, sex effects in respiratory system variables are minor and can largely be attributed to differences in body size and/or intensity levels (Hoit et al., 1990; Polgar & Weng, 1979; Stathopoulos & Sapienza, 1997). Subglottal pressure and patterns of respiratory volume change during speech do not appear to vary with sex (Holmberg, Hillman, & Perkell, 1988; Netsell et al., 1994; Russell & Stathopoulos, 1988; Stathopoulos & Sapienza, 1993a; Trullinger & Emanuel, 1983). Sex effects in some airflow measures probably reflect differences in laryngeal structure. For example, higher AC airflows, or pulse amplitudes, have been found in men than in women or children (Löfqvist, Koenig, & McGowan, 1995; Stathopoulos & Sapienza, 1993b), as expected given that men's longer vocal folds permit greater lateral amplitudes of vibration. Average DC flow rates appear to be higher in men than women during normal expiration (Polgar & Weng, 1979) and in some speech contexts (cf. Stathopoulos & Weismer, 1985a, 1985b), and are lowest in children. These effects on average flow rates are usually attributed to differences in airway resistance: Children, with the smallest average cross-sectional areas, have the highest airway resistances, whereas men have the lowest resistances, and women are intermediate (Holmberg et al., 1988; Stathopoulos & Weismer, 1985a, 1985b; Trullinger & Emanuel, 1983).

Use of /h/ to Study Group Differences in Laryngeal and Aerodynamic Quantities

The observations above suggest that systematic group variation exists in many or all of the factors directly involved in voicing control. Indeed, a sizable body of literature, particularly on voice quality, indicates that vibratory characteristics do vary across populations in ways that are consistent with these physical differences. For example, it has been established that breathy voice qualities are more often observed in women than men (e.g., Fant, 1993; Holmberg et al., 1988; Klatt & Klatt, 1990; Monsen & Engebretson, 1977; Södersten & Lindestad, 1990), as expected if women have a greater prevalence of posterior glottal chinks during phonation. Voice quality data are an imperfect source of information on laryngeal structure and function, however. Speakers have considerable latitude for selecting voice qualities, and their choices could reflect social and/or cultural norms (cf. Bickley & Stevens, 1986; Hanson, 1997; Klatt & Klatt, 1990; Södersten & Lindestad, 1990), so group differences may indicate behavioral as well as anatomical or physiological differences.

The consonant /h/ offers another method of assessing laryngeal control and voicing behavior in running speech. Because /h/ has a fairly open vocal tract and involves little or no upper articulator movement, variations in glottal area (as for abduction) can be monitored noninvasively by means of an oral airflow signal. Using the flow signal as a measure of abduction assumes that laryngeal resistance varies mainly with gross changes in cross-sectional area and that subglottal pressure remains essentially constant. In fact, subglottal pressure may undergo a dip of 1–3 cm H_20 in the vicinity of a major abduction gesture (Löfqvist, 1975; Ohala, 1990; Ohala & Ohala, 1972; Slis & Damsté, 1967). A decrease in pressure should diminish the rate of flow increase for large glottal areas, so that oral flow signals may underestimate the extent of glottal area changes. Accordingly, ranges of flow change during /h/, within and across individuals, should be regarded as conservative estimates of variation in abduction degree.

In English and many other languages, /h/ has been traditionally categorized as voiceless, but descriptive phoneticians have long recognized that voiced variants occur frequently in natural speech, especially between voiced segments (e.g., vowels) and in unstressed positions, where abduction extent may be limited (Fant, 1993; Gobl, 1988; Pierrehumbert & Talkin, 1992; Rothenberg & Mahshie, 1988). Because voicing is noncontrastive in English /h/, a period of voicelessness per se is not mandatory for accurate identification. In some speakers, the acoustic manifestation of /h/ may be just a region of breathy voicing (Manuel & Stevens, 1989). Thus, the production requirement appears to be simply some degree of abduction, or some change in glottal configuration that produces characteristics associated with breathiness, such as sinusoidal pulse shapes, higher open quotients, weaker vocal tract excitations, higher spectral tilt, and increased DC airflow (Fant, 1993; Gobl, 1988; Pierrehumbert & Talkin, 1992; Rothenberg & Mahshie, 1988). A period of breathiness is often evident as speakers move into and out of abducted postures (Fant, 1993, 1995; Gobl, 1988; Gobl & Ní Chasaide, 1988; Klatt, Stevens, & Mead, 1968; Lindqvist, 1972a; Löfqvist et al., 1995; Rothenberg, 1972; cf. also Hertegård, Gauffin, & Lindestad, 1995). Recent work on voice quality suggests that speakers may also be able to achieve some characteristics of breathy voicing by making laryngeal adjustments that alter vibratory properties but do not necessarily involve much change in glottal area (Hanson, 1997).

The fact that /h/ has voiced and voiceless variants makes it uniquely useful for studying the laryngeal conditions that affect voicing. Speakers presumably do not need to exercise the same degree of direct control over laryngeal settings for /h/ as for contrastively voiced consonants, allowing the voicing status of /h/ to follow straightforwardly from current glottal conditions. Also, unlike voice quality, /h/ voicing is not known to carry significant pragmatic or affective information. Hence, group differences are more directly attributable to physical factors such as laryngeal structure and/or aerodynamic quantities. Consistent differences in /h/ voicing across groups would suggest that a characteristic abduction gesture yields different results, depending on the speaker's physical and structural characteristics. This, in turn, implies that speaker groups may need to adopt different strategies for achieving consonantal voicing contrasts. The present study specifically investigates how /h/ voicing may vary among men, women, and children, and how voicing of /h/ is related to voicing patterns in oral stop consonants.

Method

Participants

Data were collected from 7 speakers in each of three groups: adult males (participants AM1–7, age range = 26.6–57.7 years, mean = 35.4 years), adult females (AF1–7, age range = 27.1–51.3 years, mean = 37.5 years), and 5-year-old children (girls 5F1–3 and boys 5M1–4, age range = 4.7–5.9 years, mean = 5.4 years). Individuals are identified hereafter by their group code and number; for example, AF3 is the 3rd participant in the adult female group. All participants were normal, healthy, native speakers of American English. Parents of the 5-year-olds filled out questionnaires about the child's developmental history and were questioned to ascertain that there was no known or suspected history of speech, language, hearing, or developmental disorders and no extensive exposure to languages other than English. A sample of each child's spontaneous speech was recorded on tape and analyzed using Miller and Chapman's (1993) Systematic Analysis of Language Transcripts (SALT) programs. Results confirmed that all children showed a mean length of utterance (MLU) within or above expected age ranges, according to Miller and Chapman (1981). A certified speech-language pathologist subsequently reviewed the tapes and verified that the children showed no signs of articulatory disorders.

Speech Materials and Elicitation

Speech materials were 4-syllable nonsense utterances of the form [ˌmɑmə ˈC̲ɑpə], in which the third syllable received primary stress and began with one of the consonants /b d p t h/. The form of this utterance was chosen to place the target syllable in a running speech context that 5-year-olds could easily learn and remember, and to avoid laryngeal adjustments or intraoral pressure changes in the syllable preceding the target. Each utterance was presented to participants verbally, beginning with [ˌmɑmə ˈpʰɑpə], the one most easily identified as an ordinary English utterance. The participant then produced multiple repetitions of this at a self-selected rate for one or more trials until many tokens of the utterance had been collected. Adults usually produced 10–15 repetitions per trial, with 2–4 trials per consonant, whereas child participants usually produced 4–8 repetitions per trial, with 4 or more trials per consonant. More tokens were collected of /p t h/ than of /b d/ because the voiceless consonants were of primary interest. The number of tokens we could reasonably expect to gather from our child participants was 20–30 per voiceless consonant. More tokens were collected from the adults to increase the likelihood that subtle gender differences might be detected. No participant objected to any of these novel utterances or showed any signs of difficulty with the task.

Recording Procedures and Equipment

For every utterance, airflow was recorded using one of three differently sized Rothenberg masks (Glottal Enterprises models MA-1N, MA-1S, and MA-1P). Participants were carefully instructed to keep the mask pressed firmly into their faces during recording so as to prevent air leaks. Adults held the mask themselves during recordings; for child participants, a second experimenter either held the mask or else monitored the child closely to make sure the mask was pressed tightly to the face. The masks used here cover both the nose and the mouth, so the output signal represents oral and nasal flow combined. Nasal airflow was assumed to be negligible during the oral portions of the utterance. Some nasalization of vowels was observed in the [ˌmɑmə] portion of the carrier phrase, as evidenced by lower amplitudes and simpler waveshapes (reduced high-frequency content) in the airflow signals. In the target syllable, however, the vowel was always flanked by oral consonants, and no such evidence of nasalization was seen here for any speaker. Acoustic signals were also collected for every utterance, using a standing microphone positioned near the participant. Because the acoustic signal received outside the mask is somewhat damped, this was not used for measurement purposes, but primarily to verify the accuracy of the experimental utterances. Finally, for a subset of utterances containing /h p b/, an intraoral air pressure signal was recorded via an open-ended flexible tube that passed through one of the orifices in the mask to rest between the participant's lips during bilabial closure at a roughly horizontal angle. A catheter-tip pressure transducer was fed into the tube and positioned with its tip just inside the mask.

Acoustic signals were filtered at 9.5 kHz and digitized at 20 kHz, and pressure and flow signals were filtered at 4.5 kHz and digitized at 10 kHz, all with 12-bit precision. Immediately after each recording session, flow calibration signals were obtained using a rotameter, and pressure calibration signals were collected using a water manometer.

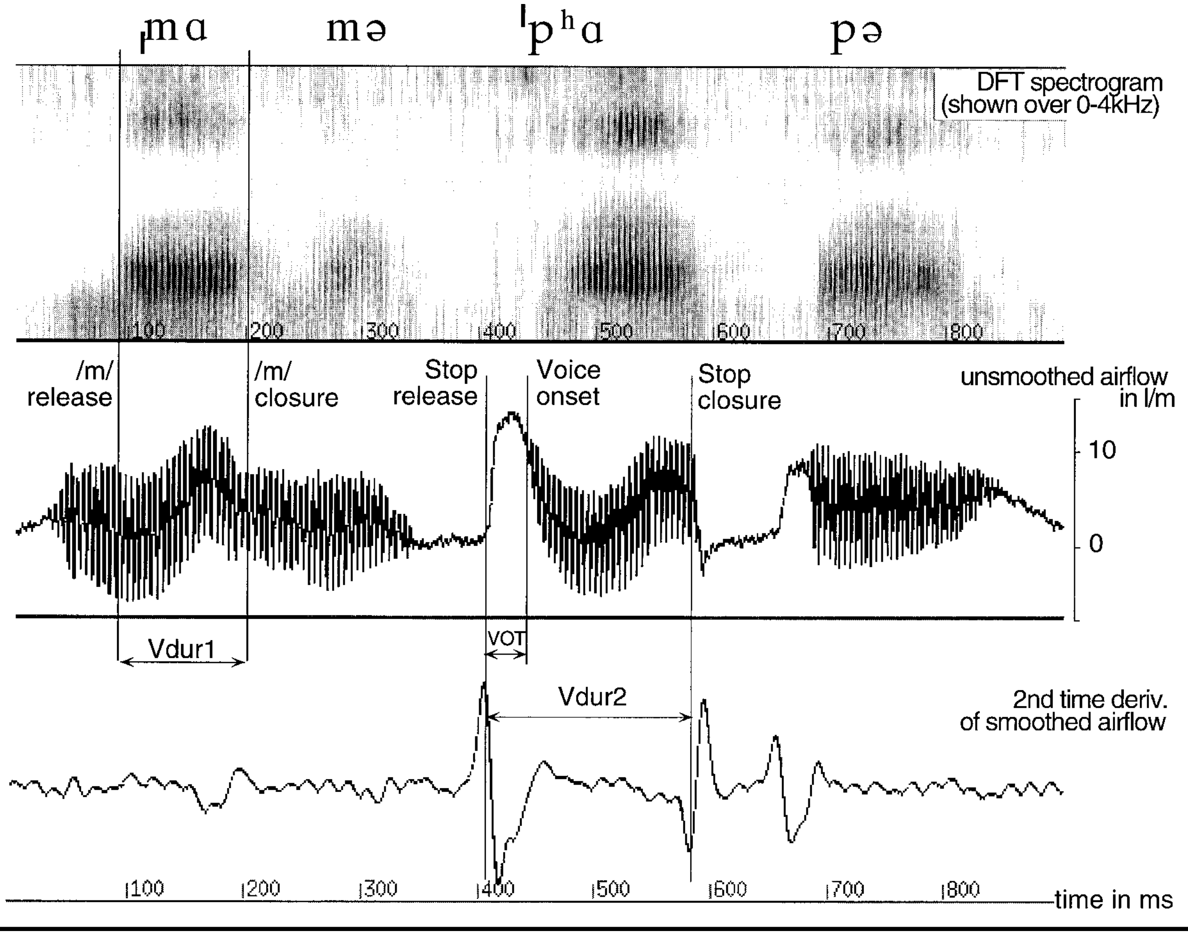
Signal Processing

Signals were stored on a VAX computer and processed using the HADES signal analysis program at Haskins Laboratories (Rubin & Löfqvist, 1996). The flow and pressure signals were smoothed with a wide triangular window (133 points) to obliterate all or most evidence of glottal pulses. From these, the first time derivatives of the flow and pressure signals were obtained, using a 3-point difference algorithm, and smoothed with a 133-point window. For utterances containing target stop consonants, the second time derivative was also obtained and smoothed. Finally, the original flow signals were smoothed twice, consecutively, with a narrow triangular window (5 points) to eliminate noise and allow easier identification of glottal pulses.

Measurement Procedures and Criteria

The two events that define VOT, namely stop release and voicing onset, were located in the airflow signal as shown in Figure 1. The stop releases in the target syllable were defined according to the local peak in the second time derivative of the smoothed flow signal, representing the rapid airflow increase that occurs upon oral release. Voicing onset was set at the first visible pulse in the original (lightly smoothed) flow signal, and VOT was calculated as the difference between these two times. The potential measurement error for visually defined voicing onset is estimated to be one glottal period so is, in general, highest for the adult male speakers with the lowest fundamental frequencies. For /b d/, voicing from the preceding vowel sometimes continued uninterrupted through the consonantal closure. Traditionally, VOT is not measured in such cases because voicing onset for the stop cannot be measured as an

Figure 1. Labels used to define stop VOTs and vowel duration. The /m/ releases and closures were determined according to changes in amplitude and higher frequency components in the airflow and, secondarily, the acoustic signal. The /p/ releases and closures were set at peaks and valleys (respectively) in the second time derivative of the smoothed airflow signal. Voice onset was determined visually from the unsmoothed flow signal.

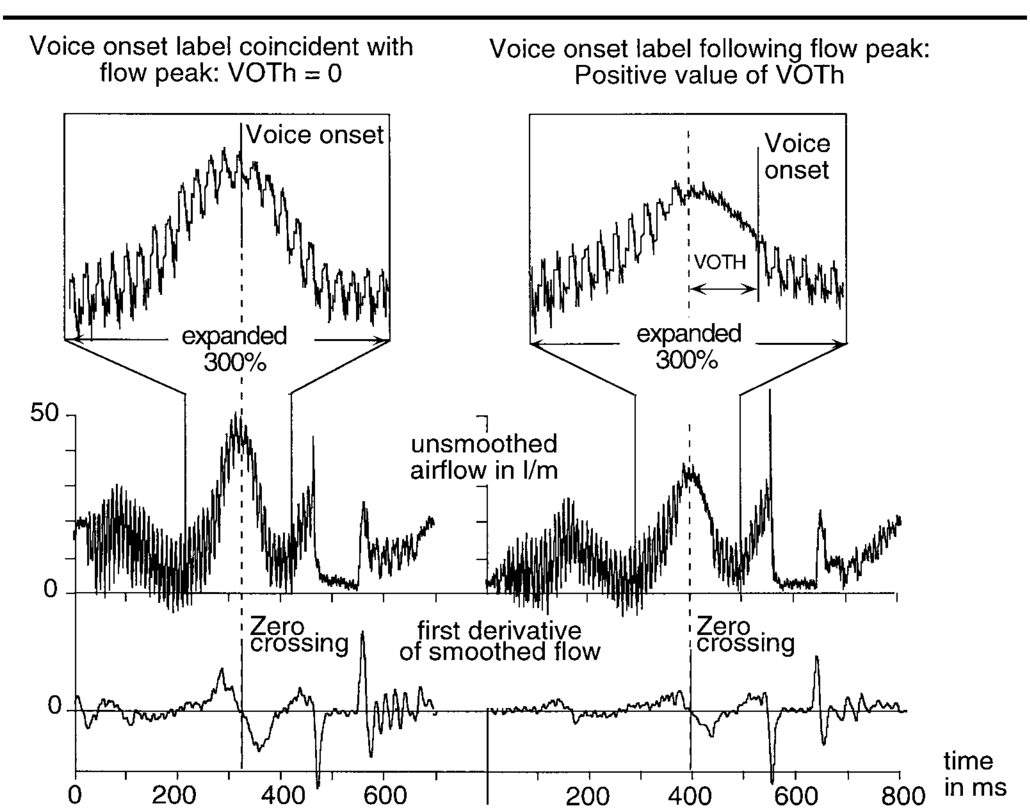


independent event (cf. Lisker & Abramson, 1967). Because the current study involved children, however, it was desirable to assign a numerical value of VOT to all stops in order to perform statistical testing of category contrast in /b p/ and /d t/ pairs. Therefore, tokens of /b d/ that showed continuous voicing through the closure were assigned a VOT value of 0 ms. This value was chosen to yield the most conservative estimate of a participant's productive VOT contrast; in other words, participants who produced fully voiced stops actually had greater phonetic differentiation of the two voicing categories than the statistics reflect.

Two measures of vowel duration were made to obtain an estimate of participants' speaking rates. To keep the measurement task manageable, these were made during utterances containing target /p/ only. Both measured an interval from consonantal release to consonantal closure. The first vowel in [ˌmɑmə] (labeled "Vdur1" in Figure 1) was measured from the release of the initial /m/ to the closure of the second /m/. Release and closure were defined as the moments where the flow signal showed sudden changes in amplitude and/or high-frequency components as compared to the intervening vowel. Appearance of higher frequency components in a wide-band spectrogram of the acoustic signal was sometimes a useful additional indication of /m/ release and closure when the vowel was somewhat nasalized. The second vowel duration ("Vdur2") was measured in the stressed syllable, from the release of the target /p/ to the closure of the /p/ for the final syllable. The closure was defined according to the local minimum in the second time derivative of the flow signal. Some of the children (5M4, 5F2, 5F3) regularly produced the first vowel with extensive nasalization, such that there were no clear pulse shape or amplitude changes differentiating consonantal and vocalic regions. One child (5F3) frequently devoiced the end of the stressed vowel. These factors precluded making vowel duration measures for these participants, so summary vowel duration data are not reported for them.

Figure 2 shows the labels used to define flow peaks and voicing onsets in /h/. Flow peaks were set at times of major zero crossings in the first time derivative of the smoothed signal. The peak flow amplitudes in /h/ serve as an indication of abduction extent. The voicing characteristics of /h/ were measured by calculating the time of voice onset relative to the flow peak. Insofar as this measure quantifies when voicing begins relative to

Figure 2. Labeling VOTh for tokens with and without voicing breaks. The flow peak was set at the zero crossing in the first time derivative of the smoothed airflow signal. The time between flow peak and voice onset was defined as VOTh.



another event (here, the flow peak rather than plosive release), it is referred to as VOTh. The duration measured by VOTh is intended to be analogous to the VOT in a voiceless aspirated stop, where peak glottal abduction occurs nearly coincident with oral release. Cases of unbroken voicing throughout the /h/ were defined as having VOTh = 0, whereas a cessation of voicing around the /h/ flow peak yielded a positive value of VOTh. In some participants, noise and aperiodicity around the /h/ flow peak complicated visual identification of low-amplitude glottal pulses around voicing onsets. Consequently, voicing onsets for /h/ were defined from spectral changes in the unsmoothed flow signals. A narrow-band Discrete Fourier Transform was used to create a waterfall display of the region around the flow peak (DFT window size = 512 points, Hamming window, with a slide of 64 points and overlap of 448 points between adjacent windows). By stepping through each token one frame at a time (steps of about 6.4 ms), voicing onsets could be seen as abrupt spectral changes that typically included emergence of the first and second harmonics, and sometimes higher harmonics as well. To verify that this method yielded similar results to visually defined voicing onsets, VOTh was measured in several participants using both methods (4 women, 2 men, and 3 children, for a total of 335 tokens of /h/). The VOTh values obtained by the two methods were highly correlated (r = .968, p < .0001), with an average durational difference of less than 2 ms. Finally, pressure peaks during the unstressed /p/ in the last syllable of the utterance were defined using the first time derivative of the smoothed pressure signal, in a manner analogous to that described for the /h/ flow peaks above. The peak pressure values provide an estimate of the speakers' subglottal pressure levels.

Once all measurement points were marked with labeled cursors, software procedures in HADES were used to obtain time and amplitude values at those positions. The raw data were subjected to distributional analysis using the BMDP software package (Release 7), and the summaries were transferred to a Macintosh computer for tabulation, plotting, and statistical treatment using StatView (version 4.5).

To provide an index of measurement reliability, 5% of all tokens were reanalyzed by the same experimenter (the author), using identical procedures, several months after the original measurements were made. Correlational analyses of the two sets of measurements showed near-perfect agreement for the pressure (r > .9995, p < .0001) and flow peak (r = .999, p < .0001) values. For VOT, VOTh, and vowel duration measures, all correlations were r > .96 and p < .0001, with an average durational difference less than 2 ms.

Results

Tables 1–4 give individual and group averages for stop VOTs, vowel durations, aerodynamic measures, and VOTh. Because /p t/ VOTs in many participants showed significant positive (rightward) skew, Mann-Whitney U tests were used instead of the t test to establish significant category contrast, and medians as well as means were entered into the analyses of variance (ANOVAs) used to test for differences between groups. To test for group differences in variability, standard deviations (SDs) and coefficients of variation ($CoV = SD/M$) were also entered into the ANOVAs. The coefficient of variation was included to correct for cases where higher standard deviations might arise simply as a function of longer durations. Extreme departure from normality in the /h/ voicing data precluded the use of central tendency measures. In particular, participants who produced predominantly, but not exclusively, voiced /h/ showed mean and median values very close to zero, leaving these participants virtually indistinguishable from those who always produced fully voiced /h/. Accordingly, analysis of /h/ relied on individuals' percentages of voiced /h/ (%VOTh = 0) and ranges of VOTh. Finally, for many participants, the voiced and devoiced tokens of /h/ formed distinct distributions, so VOTh ranges were calculated both with and without cases of VOTh = 0. The ANOVA results for all variables are summarized in Table 5. Results with $p \leq .05$ were defined as being statistically significant.

Stop VOTs and Vowel Duration Measures

The U tests indicated that all children and adults showed significant VOT differences within /p b/ and /t d/ pairs. No significant differences were found among the three groups for means or medians of the stop VOTs (Table 1) or the stressed vowel durations (Table 2). The children did show significantly longer unstressed vowel durations than either adult group, but this result must be treated with caution because data were only obtained for 4 of the 7 children. The clearest group differences emerged in the variability measures. The 5-year-olds had higher standard deviations than adults on /p t/ VOTs and vowel durations. Results for the coefficient of variation were similar to those for the standard deviation, though significance levels were diminished for the unstressed vowel comparison.

Aerodynamic Measures

The peak pressure and flow values are given in Table 3. The ANOVA showed no significant differences in mean or median pressure values for the three participant

Table 1. Distributional statistics for VOTs (all durations in ms).

Men		AM1	AM2	AM3	AM4	AM5	AM6	AM7	Men: Avg
/b/ VOT	*n*	59	34	24	15	24	20	16	192
	M	14.8	12.2	11.9	4.0	6.7	9.0	13.1	10.2
	SD	3.2	3.1	4.4	5.4	2.8	4.5	6.3	4.2
/p/ VOT	*n*	53	71	60	58	83	48	40	413
	M	56.7	41.8	46.9	34.3	44.3	38.0	36.4	42.6
	SD	15.1	14.1	14.6	7.0	11.5	6.8	7.2	10.9
/d/ VOT	*n*	12	24	12	15	24	9	16	112
	M	14.6	10.6	12.9	11.7	6.3	14.4	12.5	11.9
	SD	7.5	7.0	4.0	8.6	4.5	6.8	4.8	6.2
/t/ VOT	*n*	13	48	36	30	48	40	32	247
	M	66.5	43.4	62.2	43.8	53.3	43.5	35.9	49.8
	SD	10.5	10.6	9.8	7.5	10.0	10.2	10.5	9.9
Women		**AF1**	**AF2**	**AF3**	**AF4**	**AF5**	**AF6**	**AF7**	**Women: Avg**
/b/ VOT	*n*	36	12	23	22	20	19	27	159
	M	10.8	12.9	21.5	20.2	9.3	3.4	9.3	12.5
	SD	2.5	5.4	5.9	4.8	6.5	2.9	5.1	4.7
/p/ VOT	*n*	36	36	55	72	49	57	79	384
	M	43.6	48.5	44.5	47.2	36.6	36.2	79.7	48.1
	SD	10.9	10.9	11.3	14.0	9.3	11.8	15.5	12.0
/d/ VOT	*n*	13	12	24	12	10	28	15	114
	M	13.8	13.3	17.9	18.8	11.5	19.5	8.0	14.7
	SD	3.0	2.5	5.3	4.8	8.5	7.1	5.3	5.2
/t/ VOT	*n*	24	24	41	36	10	32	29	196
	M	47.3	64.8	51.3	68.3	65.0	68.8	66.0	61.6
	SD	7.9	7.4	12.2	13.2	10.0	13.2	10.9	10.7
5-year-olds		**5M1**	**5M2**	**5M3**	**5M4**	**5F1**	**5F2**	**5F3**	**5-year-olds: Avg**
/b/ VOT	*n*	12	10	12	12	12	5	13	76
	M	21.7	15.0	11.7	7.5	16.3	7.0	8.8	12.6
	SD	12.9	2.4	6.2	4.5	5.3	5.7	5.8	6.1
/p/ VOT	*n*	55	36	20	30	27	18	30	216
	M	104.0	38.8	27.0	36.3	58.9	55.3	48.0	52.6
	SD	24.6	21.9	11.6	17.3	28.7	14.8	27.6	20.9
/d/ VOT	*n*	10	11	8	15	12	7	11	74
	M	8.0	12.3	10.6	6.7	16.3	17.1	12.7	12.0
	SD	5.9	6.5	7.3	4.1	4.8	5.7	6.8	5.9
/t/ VOT	*n*	20	22	22	25	24	21	22	156
	M	107.8	54.5	38.9	46.0	60.2	47.4	52.7	58.2
	SD	32.2	23.0	22.0	22.8	23.7	13.1	15.9	21.8

groups. The ANOVA results on the /h/ flow peak data were close to significance ($p = .054$), and post hoc Scheffé comparisons showed that the difference between the men and children also approached significance ($p = .057$). Variability measures for the pressure and flow data were generally higher for the children than the adults. For the pressure data, the effect was significant for standard deviations and coefficients of variation, and post hoc tests showed significant differences between the men and children on both measures. For the flow peak data, the standard deviations showed no significant differences among the three groups, but coefficients of variation did, and post hoc tests again showed a significant difference between men and children.

VOTh and Comparison of Voicing in /h p t/

Clear group differences emerged for the /h/ voicing data (Table 4). The men typically produced /h/ with unbroken voicing; percentages of VOTh = 0 for individuals ranged from 60–100%. All 7 children produced a

Table 2. Distributional statistics for vowel durations (all durations in ms).

Men		AM1	AM2	AM3	AM4	AM5	AM6	AM7	Men: Avg
Vdur1	*n*	50	58	60	58	83	47	40	396
(unstr)	*M*	192.8	88.9	88.6	90.7	111.7	107.6	91.6	110.3
	SD	37.9	18.4	14.4	14.3	12.4	19.0	10.7	18.2
Vdur2	*n*	50	58	60	58	83	47	40	396
(stressed)	*M*	223.7	187.5	161.7	169.1	184.8	146.5	142.2	173.6
	SD	22.3	13.2	15.0	11.2	10.5	14.7	17.1	14.8

Women		AF1	AF2	AF3	AF4	AF5	AF6	AF7	Women: Avg
Vdur1	*n*	60	60	59	62	48	57	74	420
(unstr)	*M*	102.3	106.8	103.7	85.5	104.6	116.1	103.1	103.2
	SD	16.4	10.1	9.8	11.7	13.0	22.9	15.3	14.2
Vdur2	*n*	60	60	59	62	48	57	74	420
(stressed)	*M*	146.3	195.7	180.5	165.7	170.9	172.9	203.9	176.6
	SD	19.0	11.5	12.0	11.7	11.7	26.3	17.3	15.6

5-year-olds		5M1	5M2	5M3	5M4	5F1	5F2	5F3	5-year-olds: Avg
Vdur1	*n*	54	34	19	0	27	0	0	134
(unstr)	*M*	343.0	218.5	164.7		121.1			211.8
	SD	62.1	45.1	46.1		26.8			45.0
Vdur2	*n*	54	34	19	0	27	18	34	186
(stressed)	*M*	383.8	215.8	217.0		204.7	149.5	126.9	216.3
	SD	72.9	50.2	60.0		28.0	22.0	18.1	41.9

Table 3. Distributional statistics for pressures and flows.

Men		AM1	AM2	AM3	AM4	AM5	AM6	AM7	Men: Avg
Pres in /p/	*n*	0	130	132	73	119	98	80	632
in cm H_20	*M*		5.0	7.1	6.6	8.6	4.6	4.6	6.1
	SD		0.7	1.7	0.5	0.5	0.4	0.6	0.7
Pkflow on /h/	*n*	48	33	48	30	48	40	32	279
in l/m	*M*	63.1	50.3	44.8	35.0	33.0	48.5	33.0	44.0
	SD	13.9	7.4	10.9	4.2	5.2	5.5	4.2	7.3

Women		AF1	AF2	AF3	AF4	AF5	AF6	AF7	Women: Avg
Pres in /p/	*n*	72	60	111	118	59	101	0	521
in cm H_20	*M*	4.9	6.7	5.3	4.2	9.4	7.3		6.3
	SD	0.8	0.6	0.4	0.4	0.8	1.0		0.6
Pkflow on /h/	*n*	36	24	44	48	40	48	53	293
in l/m	*M*	68.3	64.4	15.3	11.8	31.8	3.1	46.8	34.5
	SD	19.5	9.1	3.9	2.8	8.7	3.0	6.7	7.7

5-year-olds		5M1	5M2	5M3	5M4	5F1	5F2	5F3	5-year-olds: Avg
Pres in /p/	*n*	77	43	0	45	53	23	32	273
in cm H_20	*M*	4.3	5.2		9.2	7.0	7.8	6.8	6.7
	SD	1.5	1.2		1.7	0.8	1.4	1.1	1.3
Pkflow on /h/	*n*	38	22	23	41	30	12	42	208
in l/m	*M*	12.2	30.9	23.7	23.5	5.8	8.3	25.1	18.5
	SD	4.9	17.2	10.9	9.7	6.3	3.3	9.3	8.8

Table 4. Distributional statistics for VOTh (all durations in ms).

Men		AM1	AM2	AM3	AM4	AM5	AM6	AM7	Men: Avg
VOTh	*n*	48	33	48	30	48	40	32	279
Range (all VOTh)		85	0	30	45	0	70	0	32.9
Range (VOTh>0)		60	n.a.	10	5	n.a.	50	n.a.	31.3
n VOTh = 0		29	33	44	28	48	25	32	239
% VOTh = 0		60%	100%	92%	93%	100%	63%	100%	87%
Women		**AF1**	**AF2**	**AF3**	**AF4**	**AF5**	**AF6**	**AF7**	**Women: Avg**
VOTh	*n*	36	24	44	48	40	46	53	291
Range (all VOTh)		65	30	50	75	60	0	90	52.9
Range (VOTh>0)		45	25	50	60	55	n.a.	90	54.2
n VOTh = 0		9	0	40	8	5	46	0	108
% VOTh = 0		25%	0%	91%	17%	13%	100%	0%	35%
5-year-olds		**5M1**	**5M2**	**5M3**	**5M4**	**5F1**	**5F2**	**5F3**	**5-year-olds: Avg**
VOTh	*n*	38	22	23	41	26	12	42	192
Range (all VOTh)		135	120	115	90	85	50	155	107.1
Range (VOTh > 0)		75	80	110	70	60	20	150	80.7
n VOTh = 0		13	12	17	38	13	8	8	109
% VOTh = 0		34%	55%	74%	93%	50%	67%	19%	56%

number of fully voiced /h/s, along with some devoiced tokens. The adult female group was the most heterogeneous: Some women had almost exclusively voiced /h/; others had few or no fully voiced tokens; and others produced a mixture of voiced and voiceless tokens. The ANOVA on the percentage of voiced /h/ revealed a significant group effect, with post hoc tests showing fewer voiced /h/s in women than men. VOTh ranges were significantly wider for the children than the adults, when all tokens of /h/ were included. When cases of voiced /h/ were removed and the ANOVA was re-run on the devoiced tokens of /h/, results were not significant (p = .075).

Figure 3 combines the histograms for /p t/ VOTs and VOTh into a single plot for each speaker. These displays demonstrate the group differences in percentages of /h/ voicing and also reveal a correspondence between VOTh and /p t/ VOTs within individuals. In particular, the speakers with the longest values of VOTh also produced the longest stop VOTs. This is most obvious for speaker 5M1, for whom an expanded horizontal scale was needed to accommodate his VOT and VOTh values.

To investigate further relationships among the measured parameters, correlational analyses were performed on medians, maxima, and standard deviations of all variables: VOTs, vowel durations, VOThs, and aerodynamic measures. Results (see Table 6) indicated significant positive relationships between individuals' maxima, medians, and standard deviations for /p t/ VOTs and VOTh. The vowel duration measures were highly correlated with the stop VOTs measures, but less closely related to VOTh: The r values for VOTh were lower than for the stops, and the results were not significant for median values.

Because the most extreme VOT and VOTh values belonged to the 5-year-olds, the correlations between VOTh and /p t/ VOTs were rerun for the adult participants only to see if any effect remained. For maximum values, the results approached significance for /h-p/ (r = .523, p = .054) but were not significant for /h-t/ (r = .403, p = .157). For the standard deviations, results were not significant (adults-only for /h-p/, r = .203, p > .49; /h-t/, r = .002, p > .99). Inspection of Table 6 shows significant correlations among many standard deviation measures. This primarily reflects the fact that the 5-year-olds were more variable than the adults on all measures.

Characteristics of Abduction in /h/

Figure 4 plots the smoothed airflow signals for each participant's /h/ productions. Tokens were aligned at the point of maximum flow, which is at 150 ms in these plots. As discussed above, increases in oral flow for /h/ reflect the abduction. Judging from these data, all adult males used substantial glottal opening for /h/, showing flow increases of 20 l/m or more. Among the women, some had flow increases comparable to the men's, whereas others were more limited. The children's maximum flow peaks did not approach the adult maxima, but they were well within the adult range as a whole.

Table 5. Results of ANOVAs and post hoc tests.

Measures		ANOVA results			Scheffé Results		
		DF	*F*	*p*	5:F	5:M	F:M
Medians	/b/	2,18	0.301	.744			
	/d/	2,18	0.668	.525			
	/p/	2,18	0.347	.712			
	/t/	2,18	0.97	.398			
	Vdur1	2,15	6.326	.01	.016	.024	.973
	Vdur2	2,17	0.894	.428			
	Pres	2,15	0.139	.872			
	Pkflow	2,18	3.444	.054	.289	.057	.636
	VOTh	2,18	2.67	.097	.828	.294	.11
Means	/b/	2,18	0.429	.658			
	/d/	2,18	1.316	.293			
	/p/	2,18	0.573	.574			
	/t/	2,18	1.024	.379			
	Vdur1	2,15	7.067	.007	.011	.017	.965
	Vdur2	2,17	1.278	.304			
	Pres	2,15	0.211	.812			
	Pkflow	2,18	3.88	.04	.25	.042	.6
*SD*s	/b/	2,18	1.35	.284			
	/d/	2,18	0.566	.577			
	/p/	2,18	1.349	.001	.006	.002	.909
	/t/	2,18	22.251	<.001	<.001	<.001	.971
	Vdur1	2,15	15.734	<.001	<.001	.001	.725
	Vdur2	2,17	9.044	.002	.007	.006	.993
	Pres	2,15	5.394	.017	.026	.065	.886
	Pkflow	2,18	0.173	.843			
	VOTh	2,18	7.76	.004	.017	.008	.94
*CoV*s	/p/	2,18	9.728	.001	.006	.004	.990
	/t/	2,18	18.081	<.0011	<.001	<.001	.73
	Vdur1	2,15	6.288	.01	.011	.063	.567
	Vdur2	2,17	13.009	<.001	.002	.001	.983
	Pres	2,15	4.299	.033	.045	.121	.853
	Pkflow	2,18	4.656	.023	.29	.024	.385
	Ranges (all VOTh)	2,18	9.203	.002	.025	.002	.548
	Ranges (VOTh>0)	2,14	3.142	.075	.361	.082	.558
	%VOTh=0	2,18	5.263	.016	.448	.185	.017

Note. M = adult males. F = adult females. 5 = 5-year-olds. CoVs were not computed for variables where mean values could be at or close to zero. Means of VOTh were not computed given extreme non-normality of the data. Range and % = 0 were measured for VOTh only.

Table 6. Correlations among measured variables.

	p VOT	t VOT	VOTh	Pkflow	Pres	Vdur1	Vdur2
Medians							
p VOT	1.0						
t VOT	.78*	1.0					
VOTh	.64*	.65*	1.0				
Pkflow	–.10	–.20	–.01	1.0			
Pres	–.31	–.08	–.27	–.15	1.0		
Vdur1	.61*	.62*	.38	–.22	–.30	1.0	
Vdur2	.72*	.77*	.35	–.17	–.26	.89*	1.0
Maxima							
p VOT	1.0						
t VOT	.79*	1.0					
VOTh	.64*	.55*	1.0				
Pkflow	–.15	–.31	.03	1.0			
Pres	.19	.25	.04	–.04	1.0		
Vdur1	.76*	.82*	.69*	–.10	.08	1.0	
Vdur2	.72*	.88*	.46*	–.18	.06	.96*	1.0
***SD*s**							
p VOT	1.0						
t VOT	.72*	1.0					
VOTh	.64*	.72*	1.0				
Pkflow	.15	.06	.36	1.0			
Pres	.50*	.55*	.34	.28	1.0		
Vdur1	.61*	.85*	.82*	.28	.62*	1.0	
Vdur2	.46*	.89*	.71*	.21	.58*	.95*	1.0

$^*p \leq .05$

Discussion

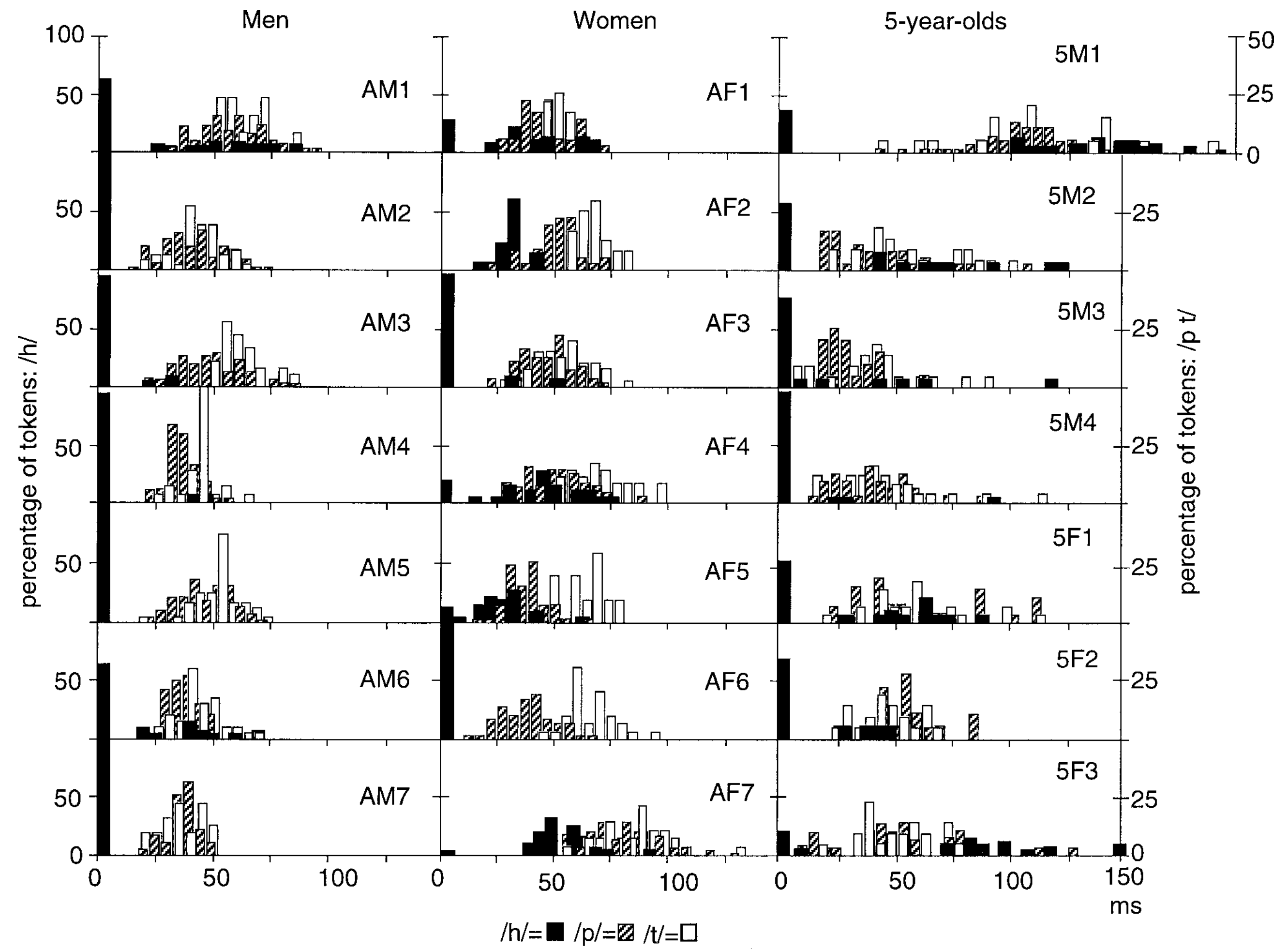
Stop VOTs

Previous reports have disagreed on when and how children's mean VOTs become adultlike. In the current data, the children were not distinguishable from the adults in terms of VOT mean or median values. Consistent with many studies, however, the 5-year-olds were more variable than adults in their productions of /p t/, as measured both by the standard deviation and the coefficient of variation (cf. Barton & Macken, 1980; Eguchi & Hirsh, 1969; Gilbert, 1977; Kent & Forner, 1980; Kewley-Port & Preston, 1974; Macken & Barton, 1980; Ohde, 1985; Ostry et al., 1984; Tyler & Saxman, 1991; Zlatin & Koenigsknecht, 1976). The fact that coefficients of variation were higher in the children implies that increased variability cannot be regarded simply as an artifact of longer durations overall (cf. Kent & Forner, 1980; Ohala, 1975; Rimac & Smith, 1984; Smith, 1978, 1994)

Aerodynamic Measures

Although group differences in pressure measures did not reach statistical significance, on average the children's pressure values were slightly higher than the adults', in keeping with previous reports (Bernthal & Beukelman, 1978; Netsell et al., 1994; Stathopoulos, 1995; Stathopoulos & Weismer, 1985b). Because the current measures were made in the last syllable of the utterance, subglottal pressures may have been diminished relative to earlier in the utterance. The values obtained here, on the order of 6–7 cm H_20, are consistent with those reported elsewhere for speakers using

Figure 3. Overlaid histograms of VOTh and /p, t/ VOTs for individual participants.



soft-to-moderate speech intensities (McAllister & Sundberg, 1996; Stathopoulos & Sapienza, 1993b).

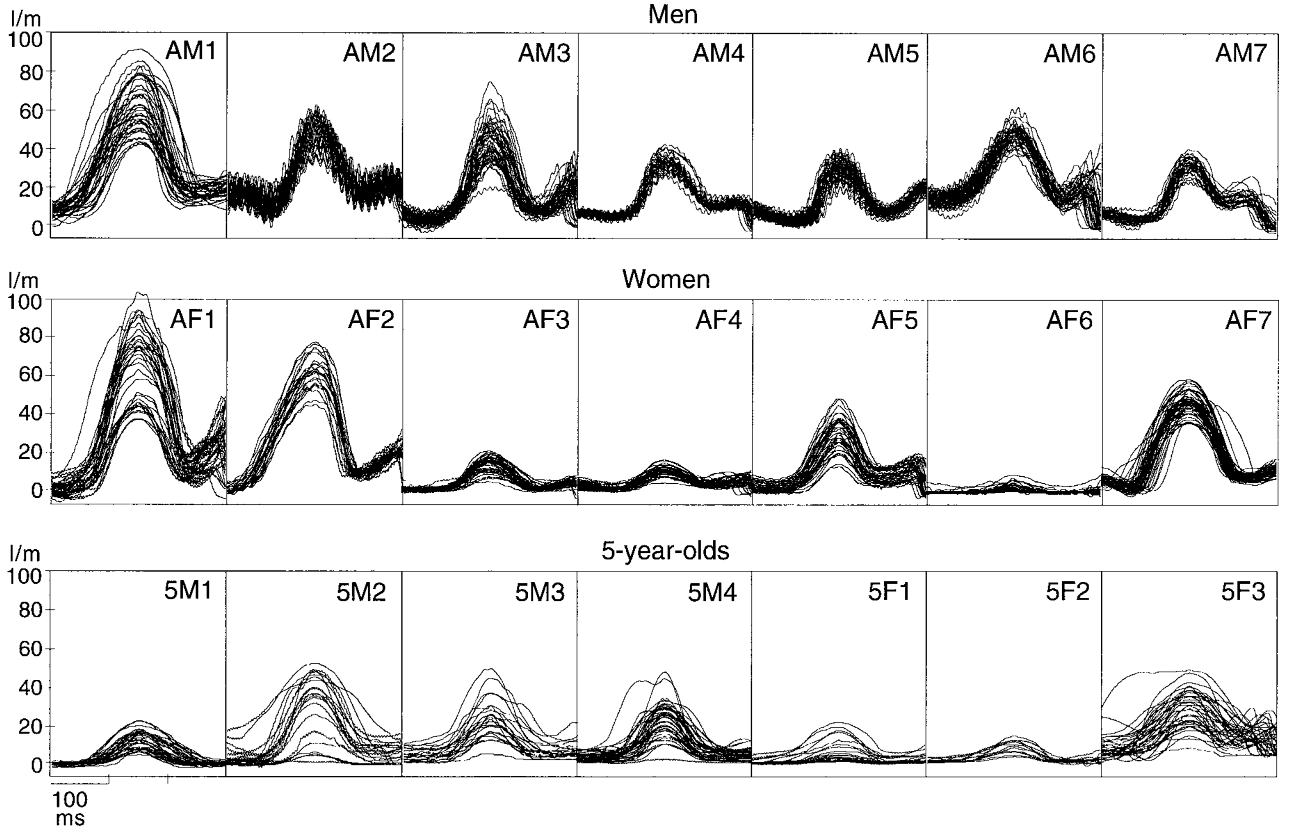
The /h/ flow peak data are consistent with previous aerodynamic studies reporting increasing flow rates with age, and higher flows in adult men than women, especially for voiceless consonants (Bernthal & Beukelman, 1978; de Troyer et al., 1978; Hoit et al., 1990; Mansell et al., 1977; Netsell et al., 1994; Polgar & Weng, 1979; Stathopoulos & Weismer, 1985a, 1985b; Trullinger & Emanuel, 1983). The men had significantly higher peak flows than the 5-year-olds, and the women's values, although not significantly different from either the men or the children, were intermediate between the two. Because airflow and subglottal pressure are directly related in an open vocal tract, one possibility was that cross-speaker variation in flow peak levels reflected differences in speaking pressure levels. To test this, a correlation was run on participants' median flow and pressure values. Results were not significant ($r = .079$, $p = .77$).

There were considerable individual differences within all participant groups in maximum flows in /h/ and in the degree of token-to-token variability. Absolute ranges of flow peak values did not differ between adults and children, nor did standard deviations. Although no attempt was made to measure the duration of /h/ flow excursions directly, perusal of Figure 4 does not suggest notable group differences on this dimension either, despite the significant differences in vowel duration noted above. The one measure of /h/ that differentiated between children and adults was the coefficient of variation. In Figure 4, this effect is evident as a wide range of peak values extending up from the bottom of a participant's flow range. In other words, the adult participants with the greatest variability were those with the highest flow peaks, whereas the 5-year-olds could show comparable variability for lower absolute measurements.

Voicing of /h/

Although voicing of /h/ has been discussed for many years in qualitative terms, this appears to be the first time it has received quantitative analysis. The VOTh

Figure 4. Oral airflow contours for multiple productions of /h/ by individual participants. Tokens were aligned at the point of maximum flow during abduction for /h/.



results support the traditional observation that voiced /h/ occurs often during running speech, but they also reveal considerable cross-speaker variation and some group trends. In an intervocalic, stressed position, fully voiced /h/ is most common among adult male speakers. In fact, many of the men here produced nothing else. Previous researchers, using data from adult male speakers, have argued that cessation of voice in running speech may not usually occur from abduction alone, but tends to require a concurrent upper vocal tract constriction (Lindqvist, 1972a, 1972b; Rothenberg & Mahshie, 1988; cf. also Sawashima, 1968). This assertion may be justified for men, but it loses validity when the participant pool is expanded to include female and child speakers. In children, voiced /h/ is common, but voiceless productions occur for many participants and may show wide ranges of VOTh values. For women, individual variation is considerable, and in some participants /h/ voicing is rare or nonexistent for the speech context considered here.

Voicing of /h/ is neither contrastive nor a known sociolinguistic variable in English. Moreover, the tendency for men to produce fully voiced /h/ apparently holds in languages other than English (viz., Lindqvist's participants were Swedish, and Sawashima's were Japanese). It therefore seems unlikely that the observed age and gender effects on VOTh distributions result from social or other speaker-selected characteristics. Rather, they probably reflect group differences in physical conditions at the glottis. In this context, it is instructive to review the aerodynamic requirements for voicing, as formalized in terms of P_{th}, the phonation threshold pressure (Titze, 1988). P_{th} is directly related to the coupling stiffness of the vocal folds, the damping of laryngeal tissues (determined by the tissue viscosity), the prephonatory glottal half-width, and the translaryngeal pressure coefficient, and inversely related to vocal fold thickness. More recent work has further indicated that phonation threshold pressure is lower for smaller glottal convergence angles and when the degree of coupling allows for a substantial vibratory phase difference between the upper and lower masses (e.g., Bickley, 1991; Lucero, 1993, 1996; Titze, 1994). Developmental changes in the larynx combine to give men longer, thicker vocal folds with lower stiffnesses than women or prepubescent children. Added tissue bulk in men may also contribute to a smaller glottal convergence angle. Increased thickness, lower stiffness, and lower convergence should all make voicing more likely for a given glottal configuration (i.e., aperture), so that we should expect men to voice under conditions where women and children do not. Conversely, women and children should devoice where men continue to voice. The VOTh data bear these expectations out. Tentatively, we may suppose that differences within and between the women and 5-year-olds also reflect variation in glottal aperture, stiffness and other tissue characteristics, and aerodynamic quantities (including subglottal pressure). These possibilities are discussed in greater detail in Koenig (1998).

Abduction for /h/

A qualitative age difference in laryngeal behavior is evident in the /h/ flow contours in Figure 4. For most of the adult speakers, one observes a clustering of tokens whereby flow excursions show consistency over both amplitude and time. In the adults who showed greater amplitude variability (particularly AM1 and AF1), one can still observe stereotypic /h/ gestures that show variation of scale, but retain a roughly common shape and rate of flow change over repetitions, so that low-amplitude tokens of /h/ tend to be short in duration whereas higher-amplitude tokens are longer. For most of the 5-year-olds, however, it would be comparatively difficult to define a "typical" abduction pattern. Tokens with high amplitude peaks are not necessarily those with the longest durations, suggesting inconsistency in rate of flow increase and decrease. The children with occasional flat or very low /h/ peaks apparently varied in whether abduction occurred at all. In contrast, the three women who produced low-amplitude /h/ peaks did so quite consistently. Finally, the children's token plots show several occurrences of multiple flow peaks or extended plateaus. Two-stage laryngeal movements have been observed in adults under direct laryngoscopy (Cooke, Ludlow, Hallett, & Selbie, 1997), and are occasionally evident in the adult signals here as well, but complex contours are much more prevalent among the 5-year-olds. Thus, it is not just the case that children are relatively more variable than adults in the degree of glottal opening they achieve. They are also more variable in how they get to their maximum abductions and in how they return to adducted postures.

The virtual absence of any flow increase for AF6 was unique among all participants studied. Post hoc inspection of her /h/ regions (determined auditorily) showed attributes of breathy voice quality, in the form of sinusoidally shaped glottal pulses and reduced amplitudes for frequencies above about 400 Hz. The minimal increases in airflow during AF6's /h/'s suggest that she may have been effecting characteristics of breathy voicing by some means other than abduction per se (cf. Hanson, 1997).

Comparison of Voicing in /h p t/

The main age effect in /h/ voicing was increased intrasubject variability among children, as evidenced by their significantly larger ranges of VOTh compared to adults. This parallels the results for /p t/, where children showed higher standard deviations and coefficients of variation than adults. As noted above, previous discussions of developmental variability in VOT for voiceless aspirated stops have referred mostly to interarticulator timing. The VOTh results indicate, however, that control of voicing varies considerably among 5-year-olds even when upper articulator movement is utterly lacking.

The /h/ flow plots and the flow peak data further suggest that 5-year-old children are not yet adult-like in their management of laryngeal factors like abduction degree and vocal fold tension, and/or in aerodynamic factors such as transglottal pressure and flow. Developing such control is presumably a prerequisite for attaining adult-like production patterns for voiceless consonants in general, including those for which voicing is contrastive. This conclusion receives support from the fact that the children with the widest VOTh ranges were the same ones who had the greatest variability in stop VOT values. Furthermore, the variability in abduction and voicing behavior for /h/ were obtained in a consonant that has no specific supraglottal configuration and which is probably less motorically challenging than a consonant with both laryngeal and supralaryngeal components.

There are, to be sure, important production differences between /h/ and the stop consonants. In particular, stop consonants involve oral obstruction, intraoral pressure increase, and a drop in transglottal pressure. Unless a speaker performs compensatory maneuvers, voicing will typically die out quickly after oral closure (Bickley & Stevens, 1986; Fant, 1995; Gobl, 1988; Lindqvist, 1972a; Ohala, 1983; Rothenberg, 1972; Rothenberg & Mahshie, 1988). The effect of this should be to make devoicing more common and/or extensive in stops than in /h/. Moreover, because /h/ voicing is not distinctive in English, speakers have no linguistic motivation for achieving a voiceless interval in /h/ as for the stop consonants. One might even expect these differences to preclude any similarities between speakers' VOTh and stop VOT distributions. The fact that /p t/ show significant correlations with /h/ at all suggests that, within an individual, some regularities hold among voiceless consonants as a class. Furthermore, the median values for the /p t/ VOTs correlated with each other, with VOTh, and with the vowel duration measures, whereas VOTh showed weaker relationships with the vowel duration measures. This too suggests that a speaker's /h/ voicing behavior reflects aspects of voiceless consonant production in particular, and not simply speech rate or other durational factors.

The men and women in this study differed significantly in their percentages of fully voiced /h/. If men have a tendency to voice over more of their abductory extents than women do, we might expect men, overall, to show slightly shorter values of VOT in voiceless aspirated stops. Although the ANOVAs on /p t/ VOTs did not show a significant gender effect, average VOTs for the women did slightly exceed those for the men (cf. Table 1). Ryalls, Zipprer, and Balduff (1997) have recently reported a significant gender effect on VOT of a magnitude similar to that seen here (about 10 ms). Further study is needed on whether adult men and women differ in their patterns of distinctive consonantal voicing. Where linguistic distinctions are involved, effects will probably be subtle, so demonstrating a reliable effect may require substantial quantities of data from many speakers. Although a fairly large number of /p/ productions was analyzed in the current study (c. 1,000), the number of participants may have been too small to reveal group differences.

Conclusions

The work reported here adds to the body of data showing that adult-like VOT distributions for voiceless aspirated stops are acquired only after a period of several years, well after a VOT category contrast may be evident in mean values. Previous accounts of VOT acquisition have tacitly assumed that children have adult-like control of the abduction gesture itself and that age differences consist merely in how that gesture is timed with respect to other events. Certainly, within an individual speaker, all else being equal, variation in interarticulator phasing should have a direct effect on VOT, and this is clearly one of the critical parameters that speakers manipulate to produce VOT contrasts. But the 5-year-old children recorded here were more variable than adults in their abduction behavior and voicing for /h/, a consonant for which voicing is noncontrastive and there is no required supraglottal articulation. Insofar as management of the larynx is needed for production of /p t/ as well as /h/, therefore, some portion of the variability in children's voiceless stop VOTs is probably best ascribed to purely laryngeal factors rather than to interarticulator timing skill. This conclusion places limits on the extent to which children's VOT data may be used as a metric of interarticulator timing control development.

In broader terms, the /h/ data indicate that a characteristic abduction gesture yields varying degrees of voice suppression among men, women, and children. If there is any consistency to an individual's abduction patterns across different speech sounds, this has general implications for the management of voicing. Population differences in the frequency of voiced /h/ suggest that the same structural and aerodynamic factors implicated in voice quality variation across age and gender may also lead to variation in consonantal voicing behavior, even where voicing is distinctive. Children and adults, and possibly also men and women, may need to adopt separate strategies for achieving consonantal voicing contrasts. The data indicate a need for broader investigation into how differences in the production mechanism may be reflected in the speech output, and the implications of such differences for management of distinctive voicing contrasts.

Acknowledgments

The work reported here was supported by NIH Grants DC-00865 and DC-00121 to Haskins Laboratories and formed a portion of a doctoral thesis submitted to the Department of Speech and Hearing Sciences at the CUNY Graduate School and University Center. I would like to thank my committee members, Katherine S. Harris, Anders Löfqvist, and Richard G. Schwartz, and my outside reader, Elaine T. Stathopoulos, for their comments and advice. Thanks also to Richard S. McGowan for his input during early phases of this work; to Len Katz and Einar Mencl for advice on statistical treatments; to Alice Faber and Melanie Campbell for assistance recording the child participants; to Sylvia Yudicé Walters for evaluating the children's recordings; to Leigh Lisker, Arthur Abramson, and Vince Gracco for consultation on a variety of theoretical and technical issues; to my participants and the parents of my child participants; and to three anonymous reviewers and Jorge C. Lucero for their comments on earlier versions of this manuscript. The /h/ voicing and flow data were presented at the 132nd meeting of the Acoustical Society of America (third joint meeting with the Acoustical Society of Japan), December 1996, Honolulu, HI. The developmental arguments were presented at the First International Conference on Voice Physiology and Biomechanics, May–June 1997, Evanston, IL.

References

Barton, D., & Macken, M. (1980). An instrumental analysis of the voicing contrast in word-initial stops in the speech of four-year-old English-speaking children. *Language and Speech*, *23*, 159–169.

Bernthal, J., & Beukelman, D. (1978). Intraoral air pressure during the production of /p/ and /b/ by children, youths, and adults. *Journal of Speech and Hearing Research*, *21*, 361–371.

Bickley, C. (1991). Vocal-fold vibration in a computer model of a larynx. In J. Gauffin & B. Hammarberg (Eds.), *Vocal fold physiology: Acoustic, perceptual, and physiological aspects of voice mechanisms* (pp. 37–46). San Diego: Singular.

Bickley, C., & Stevens, K. (1986). Effects of a vocal-tract constriction on the glottal source: Experimental and modelling studies. *Journal of Phonetics, 14*, 373–382.

Cooke, A., Ludlow, C., Hallett, N., & Selbie, W. S. (1997). Characteristics of vocal fold adduction related to voice onset. *Journal of Voice, 11*, 12–22.

Crelin, E. S. (1973). *Functional anatomy of the newborn*. New Haven: Yale University Press.

de Troyer, A., Yernault, J.-C., Englert, M., Baran, D., & Paiva, M. (1978). Evolution of intrathoracic airway mechanics during lung growth. *Journal of Applied Physiology, 44*, 521–527.

Eguchi, S., & Hirsh, I. J. (1969). Development of speech sounds in children. *Acta Otolaryngologica, Supplement 57*.

Fant, G. (1993). Some problems in voice source analysis. *Speech Communication, 13*, 7–22.

Fant, G. (1995). The L-F model revisited. *Quarterly Status and Progress Report (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm), 2–3/1995*, 119–156.

Gilbert, J. H. V. (1977). A voice onset time analysis of apical stop production in 3-year-olds. *Journal of Child Language, 4*, 103–113.

Gobl, C. (1988). Voice source dynamics in connected speech. *Quarterly Status and Progress Report (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm), 1988/1*, 123–159.

Gobl, C., & Ní Chasaide, A. (1988). The effects of adjacent voiced/voiceless consonants on the vowel voice source: A cross language study. *Quarterly Status and Progress Report (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm), 2–3/1988*, 23–59.

Goldstein, U. (1980). *An articulatory model for the vocal tracts of growing children*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Hanson, H. (1997). Glottal characteristics of female speakers: Acoustic correlates. *Journal of the Acoustical Society of America, 101*, 466–481.

Hertegård, S., Gauffin, J., & Lindestad, P.-Å. (1995). A comparison of subglottal and intraoral pressure measurements during phonation. *Journal of Voice, 9*, 149–155.

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: Contemporary research and clinical issues* (pp. 22–43). San Diego: College Hill Press.

Hoit, J., Hixon, T., Watson, P., & Morgan, W. (1990). Speech breathing in children and adolescents. *Journal of Speech and Hearing Research, 33*, 33–69.

Holmberg, E., Hillman, R., & Perkell, J. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *Journal of the Acoustical Society of America, 84*, 511–529.

Ishizaka, K., & Flanagan, J. L. (1972). Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell System Technical Journal, 51*, 1233–1268.

Kahane, J. C. (1982). Growth of the human prepubertal and pubertal larynx. *Journal of Speech and Hearing Research, 25*, 446–455.

Kent, R., & Forner, L. (1980). Speech segment durations in sentence recitations by children and adults. *Journal of Phonetics, 8*, 157–168.

Kewley-Port, D., & Preston, M. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics, 2*, 195–210.

Klatt, D., & Klatt, L. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustical Society of America, 87*, 820–857.

Klatt, D., Stevens, K., & Mead, J. (1968). Studies of articulatory activity and airflow during speech. In A. Bouyhuis (Ed.), *Sound production in man* (pp. 42–55). New York: New York Academy of Sciences.

Koenig, L. L. (1998). *An aerodynamic study of consonantal voicing control in normal men, women, and 5-year-olds*. Unpublished doctoral dissertation, CUNY Graduate School and University Center, New York, NY.

Lindqvist, J. (1972a). A descriptive model of laryngeal articulation in speech. *Quarterly Status and Progress Report (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm), 2–3/1972*, 1–9.

Lindqvist, J. (1972b). Laryngeal articulation studied on Swedish subjects. *Quarterly Status and Progress Report (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm), 2–3/1972*, 10–27.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384–422.

Lisker, L., & Abramson, A. S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech, 10*, 1–28.

Löfqvist, A. (1975). A study of subglottal pressure during the production of Swedish stops. *Journal of Phonetics, 3*, 175–189.

Löfqvist, A. (1992). Acoustic and aerodynamic effects of interarticulator timing in voiceless consonants. *Language and Speech, 35*, 15–28.

Löfqvist, A., Koenig, L. L., & McGowan, R. S. (1995). Vocal tract aerodynamics in /aCa/ utterances: Measurements. *Speech Communication, 16*, 49–66.

Lucero, J. C. (1993). Dynamics of the two-mass model of the vocal folds: Equilibria, bifurcations, and oscillation region. *Journal of the Acoustical Society of America, 94*, 3104–3111.

Lucero, J. C. (1996). Chest- and falsetto-like oscillations in a two-mass model of the vocal folds. *Journal of the Acoustical Society of America, 100*, 3355–3359.

Macken, M., & Barton, D. (1980). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language, 7*, 41–74.

Mansell, A., Bryan, A., & Levison, H. (1977). Relationship of lung recoil to lung volume and maximum expiratory flow in normal children. *Journal of Applied Physiology, 42*, 817–832.

Manuel, S., & Stevens, K. (1989). Acoustic properties of /h/. *Journal of the Acoustical Society of America, 86*, S49.

McAllister, A., & Sundberg, J. (1996). Data on subglottal pressure and SPL at varied vocal loudness and pitch in

8- to 11-year-old children. *Quarterly Status and Progress Report (Speech Transmission Laboratory, Royal Institute of Technology, Stockholm), 4/1996*, 27-36.

Menyuk, P., & Klatt, M. (1975). Voice onset time in consonant cluster production by children and adults. *Journal of Child Language*, *2*, 223–231.

Miller, J. F., & Chapman, R. S. (1981). The relation between age and mean length of utterance in morphemes. *Journal of Speech and Hearing Research*, *24*, 154–161.

Miller, J. F., & Chapman, R. S. (1993). *SALT: Systematic analysis of language transcripts*. Madison, WI: Waisman Center, University of Wisconsin.

Monsen, R. B., & Engebretson, A. M. (1977). A study of variations in the male and female glottal wave. *Journal of the Acoustical Society of America*, *62*, 981–993.

Netsell, R., Lotz, W., Peters, J., & Schulte, L. (1994). Developmental patterns of laryngeal and respiratory function for speech production. *Journal of Voice*, *8*, 123–131.

Ohala, J. (1975). The temporal regulation of speech. In G. Fant & M. Tatham (Eds.), *Auditory analysis and perception of speech* (pp. 431–453). London: Academic Press.

Ohala, J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 189–216). New York: Springer-Verlag.

Ohala, J. (1990). Respiratory activity in speech. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 23–54). Dordrecht: Kluwer.

Ohala, M., & Ohala, J. (1972). The problem of aspiration in Hindi phonetics. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo*, *6*, 39–46.

Ohde, R. N. (1985). Fundamental frequency correlates of stop consonant voicing and vowel quality in the speech of preadolescent children. *Journal of the Acoustical Society of America*, *78*, 1554–1561.

Ostry, D., Feltham, R., & Munhall, K. (1984). Characteristics of speech motor development in children. *Developmental Psychology*, *20*, 859–871.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Segment, gesture, tone* (pp. 90–117). Cambridge: Cambridge University Press.

Polgar, G., & Weng, T. R. (1979). The functional development of the respiratory system from the period of gestation to adulthood. *American Review of Respiratory Disease*, *120*, 625–695.

Rimac, R., & Smith, B. (1984). Acoustic characteristics of flap productions by American English–speaking children and adults: Implications concerning the development of speech motor control. *Journal of Phonetics*, *12*, 387–396.

Rothenberg, M. (1972). The glottal volume velocity waveform during loose and tight voiced glottal adjustments. In A. Rigault & R. Charbonneau (Eds.), *Proceedings of the 7th International Congress of Phonetic Sciences* (pp. 380–388). The Hague: Mouton.

Rothenberg, M., & Mahshie, J. J. (1988). Monitoring vocal fold abduction through vocal fold contact area. *Journal of Speech and Hearing Research*, *31*, 338–351.

Rubin, P. E. R., & Löfqvist, A. (1996). HADES (Haskins analysis display and experiment system). *Haskins Laboratories Status Report on Speech Research* (available at www.haskins.yale.edu/haskins/SR/SRweb/SRweb1.html).

Russell, N., & Stathopoulos, E. (1988). Lung volume changes in children and adults during speech production. *Journal of Speech and Hearing Research*, *31*, 146–155.

Ryalls, J., Zipprer, A., & Baldauff, P. (1997). A preliminary investigation of the effects of gender and race on voice onset time. *Journal of Speech, Language, and Hearing Research*, *40*, 642–645.

Sawashima, M. (1968). Movements of the larynx in articulation of Japanese consonants. *Annual Bulletin, Research Institute of Logopedics and Phoniatrics, University of Tokyo*, *2*, 11–20.

Sawashima, M., Abramson, A., Cooper, F., & Lisker, L. (1970). Observing laryngeal adjustments during running speech by use of a fiberoptics system. *Phonetica*, *22*, 193–201.

Slis, I. H., & Damsté, P. H. (1967). Transillumination of the glottis during voiced and voiceless consonants. *Annual Report of the Institute of Phonetics of the University of Copenhagen*, *2*, 103–109.

Smith, B. L. (1978). Temporal aspects of English speech production: A developmental perspective. *Journal of Phonetics*, *6*, 37–67.

Smith, B. L. (1994). Effects of experimental manipulations and intrinsic contrasts on relationships between duration and temporal variability in children's and adults' speech. *Journal of Phonetics*, *22*, 155–175.

Södersten, M., & Lindestad, P.-Å. (1990). Glottal closure and perceived breathiness during phonation in normally speaking subjects. *Journal of Speech and Hearing Research*, *33*, 601–611.

Stathopoulos, E. (1995). Variability revisited: An acoustic, aerodynamic, and respiratory kinematic comparison of children and adults during speech. *Journal of Phonetics*, *23*, 67–80.

Stathopoulos, E., & Sapienza, C. (1993a). Respiratory and laryngeal function of women and men during vocal intensity variation. *Journal of Speech and Hearing Research*, *36*, 64–75.

Stathopoulos, E., & Sapienza, C. (1993b). Respiratory and laryngeal measures of children during vocal intensity variation. *Journal of the Acoustical Society of America*, *94*, 2531–2543.

Stathopoulos, E., & Sapienza, C. (1997). Developmental changes in laryngeal and respiratory function with variations in sound pressure level. *Journal of Speech, Language, and Hearing Research*, *40*, 595–614.

Stathopoulos, E., & Weismer, G. (1985a). An aerodynamic study of stress in children and adults. *Journal of Phonetics*, *13*, 343–355.

Stathopoulos, E., & Weismer, G. (1985b). Oral airflow and air pressure during speech production: A comparative study of children, youths and adults. *Folia Phoniatrica*, *37*, 152–159.

Stevens, K. (1977). Physics of laryngeal behavior and larynx modes. *Phonetica*, *34*, 264–279.

Stevens, K. (1988). Modes of vocal fold vibration based on a two-section model. In O. Fujimura (Ed.), *Vocal fold* *physiology: Voice production, mechanisms and functions* (pp. 357–371). New York: Raven Press, Ltd.

Titze, I. (1988). The physics of small-amplitude oscillation of the vocal folds. *Journal of the Acoustical Society of America, 83,* 1536–1552.

Titze, I. (1989). Physiologic and acoustic differences between male and female voices. *Journal of the Acoustical Society of America, 85,* 1699–1707.

Titze, I. (1994). *Principles of voice production*. Englewood Cliffs, NJ: Prentice Hall.

Trullinger, R. W., & Emanuel, F. W. (1983). Airflow characteristics of stop-plosive consonant productions of normal-speaking children. *Journal of Speech and Hearing Research*, *26*, 202–208.

Tyler, A., & Saxman, J. (1991). Initial voicing contrast acquisition in normal and phonologically disordered children. *Applied Psycholinguistics*, *12*, 453–479.

van den Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. *Journal of Speech and Hearing Research*, *1*, 227–244.

Zlatin, M. A., & Koenigsknecht, R. A. (1976). Development of the voicing contrast: A comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research*, *19*, 93–111.

Received June 24, 1999

Accepted February 16, 2000

Contact author: Laura L. Koenig, Haskins Laboratories, 270 Crown Street, New Haven, CT, 06511. Email: koenig@haskins.yale.edu.