Air Flow Rates in Normal Speakers

ROLLAND J. VAN HATTUM, PH.D.
JOSEPH H. WORTH, M.S.
Rochester, New York

Most authorities agree that speech adequacy is strongly influenced by velopharyngeal competence (*9, 10, 14, 18*). Thus, evaluation of surgical techniques and planning for habilitative measures for persons with cleft palates have been related to judgments of velopharyngeal competence. These judgments have relied largely on observation, blowing techniques, and more recently, x-ray study, in addition to subjective appraisal of articulation, voice quality, nasal emission, and intelligibility (*1, 2, 3, 10*).

Recently, there has been an increasing effort to further objectify evaluative techniques, and refinements of methods for physiologic measurement have received significantly more experimental attention (*5, 15*). As preliminary research, prior to the study of oral air flow and air pressure patterns in the speech of individuals with cleft palate, the purposes of the present investigation are: a) to report rates of oral and nasal expelled air flow for normal speakers during certain speech activities; b) to study the variability of expelled air flow among subjects, and the variability associated with certain consonant groups; c) to compare the expelled air volumes accompanying the production of sounds in syllables with expelled air volumes accompanying sounds in connected speech samples; and d) to evaluate the adequacy of the instrumentation for recording and measurement of oral and nasal air flow during speech.

Ten adult males, ten adult females, and ten children, all judged to have normal speaking ability, performed structured speech tasks consisting of vowel-consonant-vowel combinations and short sentences. The oral and nasal air flows were detected by the use of a flowmeter, recorded, along with a synchronizing speech signal, on magnetic tape, and written out on a photographic oscillograph. These records were then marked and measured. Expelled air volume was obtained by integration of the air flow records.

Instrumentation

Of the various types of flowmeters available, a warm-wire flowmeter was chosen for two reasons: first, there was a flowmeter immediately

---

Dr. Van Hattum is associated with the State University College at Buffalo, New York. Mr. Worth is at the Eastman Dental Center, Rochester, New York. This paper was presented at the 1964 annual meeting of the American Cleft Palate Association in Los Angeles, California, and is based on research supported by PHS Research Grant DE-01837, National Institute of Dental Research.

available which had adequate sensitivity and frequency response, and which had been used successfully by others in similar research (*6*); and secondly, the warm-wire flowmeter appeared to offer the best basis for future improvement in the direction of a device less restrictive to natural speech production.

The device used here consisted of a pair of flowtronic 55A1 flowmeters. The sensitive wires of these meters were held in the two channels of an anesthetist's mask divided to capture, separately, oral flow and nasal flow, as described by Kelleher and his associates (*6*). The instruments operated on a feedback principle, insuring temperature compensation over a much wider range of temperatures than that encountered in oral and nasal air flow, and frequency response to 1000 cps, which seems adequate to detect not only articulatory changes in air flow but also the fluctuations caused by laryngeal vibration. Output was available as a meter indication of velocity (in feet per minute) and as a voltage approximately proportional to the fourth root of the velocity (0–1.0 v for 0–4000 FPM). Since each of the velocity-sensitive wires was at the center of a semicylindrical conduit of constant cross-sectional area, and since the air to be measured is considered to be constant, or practically so, in all relevant physical properties (temperature, density, relative humidity), the volume rate of air flow was directly proportional to the indicated particle velocity. The output voltages of these instruments were then approximately proportional to the fourth root of the volume rate of air flow (0–1.0 v for 0–1560 ml/sec).

The output of the two flowmeters were recorded on two channels of a Honeywell 8100 FM data tape recorder. An audio signal was recorded on a third channel. These records were written out on a Honeywell 1108 photographic oscillograph (the Visicorder). By this method, the three records could be written in synchronism, and with control of the time and amplitude scales.

Periodic comparison with available calibrated flowmeters indicated that the Flowtronic's calibration remained essentially correct. However, its response was very nonlinear, air flow being approximately proportional to the fourth power of the output voltage. Two methods might be employed to correct this nonlinearity. The output voltage might be linearized by a simple diode resistor network, or the written record might be measured with a nonlinear scale. The latter method was chosen since a relatively limited number of measurements were intended.

The mask appeared to be successful in separating oral and nasal air flow, as indicated by two procedures. First of all, one of the examiners had his nasal cavity packed with sterile gauze and the nostrils were taped so that no air could escape through the nasal cavity. A measurement was then made which showed no measureable reading on the nasal channel. Then, similarly, the mouth was taped shut and measurement revealed no significant readings for the oral channel.

Fujimura (*4*) has shown that some articulatory movements are completed in as little as 9.0 milliseconds. In order to detect any changes in flow which occur this rapidly, a system bandwidth of 400 cps was established. This allowed an audio frequency envelope to be recorded on the flow wave when phonation occurred. For measurement, the average of this envelope was determined or 'traced'.

For identification of the sounds and to re-evaluate each subject's performance, an audio tape was made on an Ampex 601 recorder. The presence of the mask made high fidelity recording impossible, but the audio record obtained was adequate for the identification of phonemic elements. As previously mentioned, the audio signal was also recorded on the data tape, and written out on the paper chart, along with the flow records. This aided the identification of the voiceless sounds.

Procedure

The structured speech materials utilized consisted of fifteen simple nonsense syllables, involving vowel-consonant-vowel combinations, and five sentences. The nonsense syllables were /ipi/, /ibi/, /imi/, /iti/, /idi/, /ini/, /isi/, /izi/, /ifi/, /ivi/, /ili/, /iri/, /iwi/, /itʃi/, and /idʒi/. The sentences were *Sonny is in the tent, See me kneel, Pete needs meat, Peter Piper picked a peck,* and *Sing a song of six-pence.* The vowel /i/ was utilized as the carrier vowel, since it would not require a great deal of mandibular excursion within the mask. Although only flow measures are being presented in this discussion, future plans call for synchronous recording, for correlation and comparison, of flow measures with pressure data. Thus, selection of experimental material had to take both measures into account.

Pitch, intensity, and duration were not controlled in spite of their relationship to flow (*13*) because it was desired to obtain as natural a sample of speech production as possible. Since Ladefoged (*7*) points out the relationship of stress to flow and pressure, each subject was trained to imitate the stress pattern of the experimenter and repeated the syllables after the experimenter. The experimenter listened to the subject's production and asked for a repetition if the stress pattern appeared to be violated.

Analysis

After each recording was made, the data were written out. These curves were 'traced' and the limits of the sounds, together with the baseline, were drawn.[1] Figure 1 shows the records of two VCV syllables, typical of the /ipi/ and /ibi/.

[1] Since the warm-wire flowmeter is inherently a rectifying device, giving positive indications for both inhalation and exhalation, indications below zero are impossible. The baseline is a horizontal line drawn on the record through the point of least deflection. If the sequence represented by the record contains at least one reversal of flow, one can be confident that the baseline, so drawn, indicates zero flow.

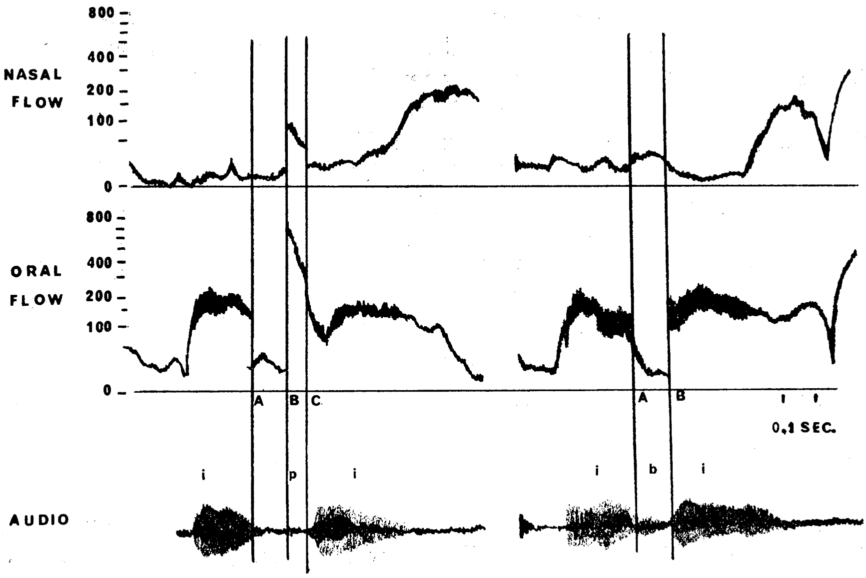

FIGURE 1. Nasal air flow, oral air flow and audio records for the two nonsense syllables /ipi/ and /ibi/, marked with baselines and consonant boundary lines A, B, and C, as described in the text. Scales are also added, showing volume rate of air flow in ml/sec. These scales are similar to the nonlinear rulers used for measuring the records.

The task of locating the consonants in the syllables and sentences proved to be difficult. The overlapping nature of the speech process made it extremely difficult to determine boundaries between vowels and consonants. Duration was found to be variable, and could not be used to provide adequate cues for locating consonants, and individual subject variation made it impossible to use characteristics of the flow curve. The /l/, /r/ and /w/ could not be identified with any degree of confidence, and these sounds were eliminated. Furthermore, the initiation and termination of the plosive phase of the /b/, /d/, and /dʒ/ could not always be identified. In addition, certain of the sounds for some of the subjects could not be identified with certainty for various other reasons and so were excluded from the data.

Included in Figure 1 are the audio traces, oral flow traces, and nasal flow traces. Also shown are the oral and nasal baselines, perpendicular lines through points A, B, and C, and the nasal and oral flow scales. Point A was determined by noting, in the oral flow trace, the first point where vocal fold activity ceased during the transition from the vowel /i/ to the voiceless consonants or, in the case of the voiced consonants, where it changed significantly in width. Magnification of records and measurement of the vibratory patterns were used where indicated to determine the points as accurately as possible. Point B was defined as the point where a periodic turbulence was first noted in the trace following the sudden expenditure of air of the explosion in the plosive sounds. Point C was taken to be the point where the explosive phase had been completed and the resumption of the vowel as evidenced by the periodic vibratory pattern was noted. Note in Figure 1 that it was not possible to determine Point C for the voiced plosives and affricate. These points were established for the nasal trace by extending the perpendiculars (through A, B, and C) from the baseline of the oral trace through the nasal trace.

The time period between points A and C was considered to be the consonant sound and duration, as well as air flow, was measured between these points. Point B was determined only for the plosive sounds and for the affricates. Distance AB was the limits of the implosive phase and distance BC was the limits of the explosive phase.

The duration of the consonant was expressed in milliseconds; oral expelled air volume and nasal expelled air volume for the consonants were expressed in milliliters. In addition, proportions were computed of oral expelled air volume to total expelled air volume and nasal expelled air volume to total expelled air volume.

The records produced by the Visicorder were crossed by 'time lines', which indicated intervals of 25 milliseconds. These lines were useful not only for determining synchronism of events on the air flow and audio traces, but also for measuring the various durations referred to above. The oral and nasal expelled air volumes were obtained by manual integration of the respective air flow records. With the aid of a specially made nonlinear ruler, air flow was measured at each intersection of an air flow curve (or of its traced average, where voiced) and a time line contained between the consonant boundary lines. Appropriate corrections were made upon terminal measurements which represented the air flow over periods which were more or less than 25 milliseconds. Each air flow measurement (ml/sec), multiplied by its effective duration (25 msec or $\frac{1}{40}$ sec), yielded the volume of air expelled during those 25 milliseconds. In practice, all the air flow measurements between a pair of boundary lines were added, and the sum multiplied by $\frac{1}{40}$ second, to yield the volume of air expelled during the interval between the two boundary lines.

Since intensity, pitch, and duration were not controlled, considerable variability in expelled air volume by subjects might be expected, as previously reported by Ladefoged (7) and Van Hattum (*17*). Such variability was indeed observed. Each of the three groups of ten speakers were found to be extremely variable and to overlap considerably in performance. Statistical differences were not found among the group scores, however, and so the 30 scores were combined into a single group.

Results

Syllables. Table 1 presents the data relating to duration and expelled air volume. In regard to duration, the unvoiced affricate /tʒ/ si the longest, followed by the unvoiced plosives /p/ and /t/ and the fricative /s/. The shortest durations were the voiced fricatives /z/ and /v/. The /b/, and /d/ and /dʒ/ were not considered because their complete duration could not be determined, since, as previously noted, the plosive phase of the consonants stage could not be differentiated clearly from the vowel which followed.

Table 1 also presents data relating to expelled air volume measures. Examination of the data relating to individual consonants is probably less meaningful than examining data for consonant groups. However, it is interesting to note that, as would be expected, the first phase of the plosives is longer in duration than the second phase but air flow is considerably less in the first than in the second phase. Note also the presence of nasally expelled air volume which was measured on the 'non-nasal' consonants.

When examining Table 2, containing mean comparisons of consonant groups, it can be seen that voicing appears to be the most important factor influencing expelled air volume, with the voiceless consonants displaying considerably more than the voiced consonants. Manner of articulation also seems to have some influence on expelled air volume, al-

TABLE 1. Means (M) and standard durations (SD) for duration (in msec), and oral, nasal, and total expelled air volume (in ml) for twelve consonants in VCV nonsense syllables, produced by 30 normal speakers. For plosive consonants, where possible values are presented for the implosive (AB) and explosive (BC) phases. BC values are not reported for the voiced plosives since an insufficient number of data were obtained to give a meaningful figure.

	Consonants																	
	/p/		/b/		/t/		/d/		/m/	/n/	/s/	/z/	/f/	/v/	/tʃ/		/dʒ/	
	AB	BC	AB	BC	AB	BC	AB	BC							AB	BC	AB	BC
Duration																		
M	137	49	120	—	127	60	106	—	119	108	186	95	173	83	121	92	108	—
SD	28	21	20		18	21	27	—	24	25	29	27	27	40	27	23	28	
Oral volume																		
M	8	25	4	—	6	30	4	—	3	2	48	17	43	14	7	32	5	—
SD	6	17	2		4	17	2		2	1	20	6	31	7	3	19	2	
Nasal volume																		
M	2	2	2	—	1	2	1	—	19	20	4	1	4	1	1	2	1	—
SD	2	1	2		1	5	1		8	10	5	1	3	0	1	4	1	
Total volume																		
M	10	28	5	—	7	31	5	—	22	22	52	18	47	15	7	35	6	—
SD	7	18	3		4	19	3		9	10	21	6	23	7	2	20	2	

TABLE 2. Comparisons of total expelled air volume (in ml) for various consonant groups, for 30 normal speakers. Values which are asterisked are based upon less than the entire population.

Difference according to voicing			
voiceless		*voiced*	
28.87		10.83*	
Differences according to manner of articulation			
plosives	*fricatives*	*affricates*	*nasals*
35.97*	32.93	44.75*	22.58
Differences according to place of articulation			
bilaterals	*labio-dentals*	*lingua-alveolars*	*lingua-palatals*
30.89	30.78	30.43*	32.41*

TABLE 3. Comparisons of means for duration (in msec) and oral, nasal, and total expelled air volume (in ml) for the /p/ in /ipi/ and in the sentence *Peter Piper picked a peck* for 20 normal adult speakers.

	/ipi/	/*p*id/	/*p*ɑɪ/	/*p*r/	/*p*ɪ kt ə/	/*p*ɛk/
Duration	177	177	158	104	158	158
Oral	34	26	29	14	27	33
Nasal	4	7	3	2	3	4
Total	38	33	32	16	30	37

though that comparison is not based on data for all subjects. (For some subjects it was not possible to measure the second phase of the plosives.) With these restrictions, it appears that the greatest volumes of air were expelled for affricates, then by the plosives, fricatives, and nasals, in decreasing amounts.

Sentences. When the production of the /p/ in the syllable /ipi/ is compared with the production of the /p/ in the sentence *Peter Piper picked a peck,* as seen in Table 3, the duration of the initial /p/ is the same as in the syllable. Three other /p/s are identical and one other, that occurring in the medial position in the word *Piper* is shorter. For oral air flow, the most noticeable difference is again the medial /p/. For nasal flow, the initial /p/ is the most different in that it displays a higher expelled air volume. However, the trace records indicate that respiration contaminates the data at the beginning and at the end of the utterance. There is frequently nasal respiration at the initiation of phonation and at the cessation of phonation. Thus, in connected speech samples, nasal air flow appears to be more affected by respiratory patterns than does oral air flow. Data relating to the /s/ in the sentence *Sing a song of six-pence* are more variable, as can be seen in Table 4. The durations of the initial and final /s/s are longer than in the syllable or within the sentence. Also, the /s/ in the word *song* is similar to what one might expect of a sound in the medial position in a word.

In general, the sounds at the beginning and at the end of utterances appear to be more variable although they tend to be longer and require more expelled air volume. However, patterns of respiration appear to contaminate measurement at the beginning and at the end of sentence utterances. Within the sentences, sounds in the medial positions in words appear to be shorter and require less expelled air volume.

Proportional Data. The proportionate amounts of oral and nasal expelled air volume were computed as percentages and are presented in Table 5. This table reveals the measured nasal air flow on so-called 'non-nasal' sounds and the measured oral air flow on the 'nasal' sounds. The standard deviations reveal the variability of some sounds and relative consistency of others. The /b/ has the most nasal expelled air volume and is the most variable, while the /s/ and /z/ are the most consistent and reveal the least nasal expelled air volume. Overall, on all oral sounds a mean of 14% nasal expelled air volume was found and on nasal sounds a mean of 12% oral expelled air volume was found.

Discussion

Voicing appeared to be the most significant determinant of the amount of expelled air volume and of the duration. Manner of articulation had some effect but place of articulation was not found to be an important variable.

In comparing the sounds in syllables and in sentences, the expelled air volume tended to be less during sentence production. The duration of the sounds studied was variable. However, it tended to be longer in the initial and final positions than in the medial position in sentences and was generally shorter in the sentences than in the syllables.

TABLE 4. Comparisons of means of consonant duration (in msec) and total expelled air volume (in ml), for the /s/ in /isi/ and in the sentence *Sing a song of sixpence*, for 20 normal adult speakers.

	/isi/	/*s*iŋ ə/	/*s*ɔŋ əv/	/*s*ɪks/	/pɛn*s*/
Duration	184	203	146	166	211
Oral	47	43	34	39	32
Nasal	4	5	4	3	6
Total	51	47	37	42	38

TABLE 5. Mean ratio of oral and nasal expelled air volume to total expelled air volume for twelve consonants in VCV nonsense syllables, according to sex and age of speaker. Since total expelled air volume is the sum of oral and nasal expelled air volumes, corresponding ratios are related, and the standard deviations shown apply to either ratio, or both. Note that SDs are the same for oral and nasal air flow rates since the values are additively related (the sum of the two is always 100).

Groups	*Consonants*											
	/p/	/b/	/m/	/t/	/d/	/n/	/s/	/z/	/f/	/v/	/ʃ/	/dʒ/
Males (N = 10)												
oral	87	60	17	92	78	15	93	94	94	90	90	83
nasal	13	40	83	8	22	85	7	6	7	10	10	17
SD	5	12	5	2	8	5	1	1	2	7	9	2
Females (N = 10)												
oral	88	70	14	88	86	10	93	93	90	92	91	87
nasal	12	30	86	12	14	90	7	7	10	8	9	13
SD	2	21	8	10	10	3	2	4	10	6	6	10
Children (N = 10)												
oral	86	77	9	91	88	10	92	91	91	93	91	92
nasal	14	23	91	9	12	90	8	9	9	7	9	8
SD	7	10	12	2	13	3	3	5	3	4	4	7
Total group (N = 30)												
oral	87	69	13	91	84	11	92	89	92	82	91	87
nasal	13	31	87	10	16	89	8	11	9	18	9	13
SD	5	16	7	7	13	5	2	5	2	6	4	9

Examination of percentage data, which reports the proportionate amounts of oral and nasal expelled air volume, reveals the presence of nasal air flow on 'non-nasal' sounds and oral air flow on 'nasal' sounds. There are several possible explanations for this. One would be that, in spite of efforts to control leakage within and around the mask, it occurred anyway.

It is possible, however, that sources other than defects of the measuring instrument might cause certain of the more-or-less anomalous air flows observed. Lubker (*8*), for instance, has hypothesized that the movement of the soft palate displaces a volume of air which would then be measured as nasal air flow. Also, movement of the closed lips inside the mask might be expected to create similar displacement-generated air flows.

Another possibility would be that air flow did occur. Van den Berg (*16*) has indicated that 'the closure of the nasal cavities by the soft palate during the production of normal non-nasal vowels is mostly incomplete'. Nusbaum, Foley, and Wells (*11*) state similar information. Rosen (*12*) reports nasal air flow on nasal sounds in normal speakers. If further research substantiates this, it could have important implications in the diagnosis of speech problems associated with cleft palates and the subsequent therapy.

The records also revealed continuous air flow on the stop consonants although not all subjects displayed such flow and, while the amount was measurable, it was nevertheless small. Again, it is very possible that faults in the instrumentation are responsible for this finding. For example, it is possible that this is really a delay in measurement of air inside the mask. This is being studied by new instrumentation (*19*) which does not have a similar problem of air storage.

Concern about the instrumentation has centered around the possible restrictions which the mask places on the articulators, the possibility of leaks between the channels and around the mask, the possibility of air storage in the mask which may result in faulty measurement, and the distance between the lips and the measuring device necessitated by the mask, again creating the possibility of faulty measurement.

In itself, the warm-wire technique appeared to function adequately; the major difficulties in the procedure were related to the mask. Questions regarding the importance of those problems must be answered by further research.

Summary

Ten adult males, ten adult females, and ten children, all normal, produced fifteen consonants in vowel-consonant-vowel syllables and in five short sentences while records of oral and nasal air flow were being obtained. Although there was considerable within-group variability, between-group differences were not significant and so all thirty subjects were combined to form a single group. Normative data for oral and nasal air flow were reported, variability among subjects, phonemes, and type of speech sample was investigated, and problems in instrumentation were discussed.

reprints: *Dr. Rolland J. Van Hattum*
1300 Elmwood
Buffalo 22, New York

Acknowledgments: This research was supported by research Grant DE-01837-01, National Institutes of Health, United States Department of Public Health. Dr. J. Daniel Subtelny is Principal Investigator. Special mention is due Dr. Joanne Subtelny for her cooperation in the collection and measurement of the data.

References

1. Buncke, H. J., Manometric evaluation of palatal function in cleft palate patients. *Plastic reconstr. Surg., 23,* 148–158, 1959.
2. Calnan, J., and Renfrew, C. E., Blowing tests and speech. *Brit. J. plastic Surg., 13,* 340–346, 1961.
3. Chase, R., An objective evaluation of palatopharyngeal competence. *Plastic reconstr. Surg., 26,* 23–29, 1960.
4. Fujimura, O., Bilabial stop and nasal consonants: A motion picture study and its acoustical implications. *J. speech hearing Res., 4,* 233–247, 1961.

5. ISSHIKI, N., and RINGEL, R., Air flow during the production of selected consonants. *J. speech hearing Res., 7,* 233–244, 1964.
6. KELLEHER, R. E., WEBSTER, R. C., COFFEY, R. J., and QUIGLEY, L. F., Nasal and oral air flow in normal and cleft palate speech: Velocity and volume studies using warm-wire flowmeter and two channel recorder (abs). *Cleft Palate Bull., 10,* 65, 1960.
7. LADEFOGED, P., Some physiological parameters in speech. *Lang. Speech, 6,* 109–119, 1963.
8. LUBKER, J., Simultaneous oral and nasal air flow and cinefluorographic measurements. Paper presented at the American Speech and Hearing Association's annual convention, 1963.
9. MOLL, K. L., Velopharyngeal closure on vowels. *J. speech hearing Res., 5,* 30–37, 1962.
10. MORRIS, H., and SMITH, JEANNE, A multiple approach for evaluating velopharyngeal competency. *J. speech hearing Dis., 3,* 218–226, 1962.
11. NUSBAUM, E. A., FOLEY, L., and WELLS, Charlotte, Experimental studies of the firmness of the velar-pharyngeal occlusion during the production of English vowels. *Speech Monog., 2,* 71–80, 1935.
12. ROSEN, C., A study of nasal emission in normal speakers. Unpublished study, The Pennsylvania State University, 1961.
13. ROUDET, L., De la depence d'air dane la parole et de ses consequence phonetique. *La Parole,* II, 21–34, 1900.
14. SHELTON, R., BANKSON, N., and BROOKS, ALTA, Nasal pressure in first grade non-cleft palate children. *J. Kansas speech hearing Assoc., 2,* 2–8, 1961.
15. SUBTELNY, JOANNE, Physio-acoustic considerations in the radiographic study of speech. *Cleft Palate J., 1,* 402–410, 1964.
16. VAN DEN BERG, JW., Modern research in experimental phonetics. *Folia Phoniat., 14,* 81–149, 1962.
17. VAN HATTUM, R., A study of air use by cleft palate and normal speaking subjects. Unpublished master's thesis, The Pennsylvania State University, May, 1952.
18. WARREN, D. W., and DUBOIS, A., A pressure-flow technique for measuring velopharyngeal orifice area during continuous speech. *Cleft Palate J., 1,* 52–71, 1964.
19. WORTH, J. H., and RUNYON, J. C., The integrating flowmeter for measuring unimpaired oral and nasal air flow. In preparation.