Journal of Voice
Vol. 6, No. 2, pp. 194–210

Special Article

International Association of Logopedics and Phoniatrics (IALP) Voice Committee Discussion of Assessment Topics

(Transcribed and edited by Diane M. Bless and Ronald J. Baken)

Discussants: Diane M. Bless, Ronald J. Baken, Thomas Hacki, Bjorn Fritzell, John Laver, Harm Schutte, Minoru Hirano, Ernst Loebell, Ingo Titze, Marie-Agnes Faure, Andre Muller, Jürgen Wendler, Sören Fex, M. Nasser Kotby, David Brewer, Aatto Sonninen, and Pertti Hurme

ELECTROGLOTTOGRAPHY

Baken: In my report I made two recommendations: that we call the vertical axis either vocal fold contact area or we call it conductance or something of that ilk. I agree with Dr. Titze that it is not possible to scale that vertical axis. There is unlikely to be a linear transform of the vocal fold contact area. I tend to support his recommendation that we not call it vocal fold contact area per se, although that term has appeared in the literature in the last several years. I do not know of a good alternative term. Conductance comes to mind for purely electronic reasons. If we call it impedance then we have to graph it upside down.

Another important point is that we cannot discern a true zero line on the electroglottogram (EGG). We can make assumptions about the beginning of contact but the so-called open phase should be flat, in theory. In theory, therefore, we should be able to say that the lowest part of the curve represents zero vocal fold contact. The fact is that we are never absolutely sure of that. The flat portion is rarely actually flat. There may be some minor vocal fold contact, perhaps at the anterior commissure, that is influencing the curve. There is also the influence, which is small but real, of other events distant from the glottis. These are periglottal events that we really cannot account for. The thing that worries me when we talk about zero lines is that many people confuse contact zero with voltage zero in their instrument. So they adjust their oscilloscope to have the zero volt line at the central line on the instrument's graticule, and they think that that means something. The fact is that it represents nothing at all. So I have tried, in instructional workshops that I have done around the United States, to avoid all mention of zero.

Sonninen: What are we really measuring? We are observing electric currents, is that not so? And we think of what the vocal folds are doing, we must interpret that from these currents. So, can we find some relevant name for the current change?

Baken: In a sense, that is what is implied by conductance.

Titze: As I said before, when there are good validation studies it is common practice to label the output of an electronic device according to what it is intended to measure. If we have a pneumotachograph or a pressure transducer, that is an electronic device and it is calibrated to read out pressure, even though it is actually outputting voltage. And that is legitimate if you have done a calibration that says "so many volts is equivalent to so many units of pressure." Unfortunately, calibration has not been done adequately enough for the EGG. Once we have done this, I would have no reservation about saying this is units of contact area. You could even put units of square centimeters to it if you have done the calibration. So we are faced with a dilemma here if we invent something too rigid in terms of a different name. Later on, we might want

Address correspondence and reprint requests to Diane M. Bless, Waisman Center, Rm. 485, University of Wisconsin-Madison, 1500 Highland Ave., Madison, WI 53705.

to change it if someone has done the necessary calibration and it does represent contact area in square centimeters.

Sonninen: Can you say what happens?

Titze: In my reporting I just label it EGG versus time. To me that is sufficient, at least for the moment, because everybody with whom I communicate knows what an EGG device is.

Baken: You and I may communicate to very different audiences, but I find that if I label it EGG or call it EGG, clinicians may have no concept at all of what EGG is. So the name conveys very little, if anything to them. And this generates massive problems in interpretation.

Titze: But, if you label it electrical conductance do they understand it any better?

Baken: No, certainly not. Typically, I have called it vocal fold contact, and that says something to the clinical audience.

Titze: That skirts the issue. What is the unit of contact?

Baken: So far, it is unitless. But remember that for clinical purposes, that might not matter. It is the rate of change or degree of change of the contact that is important. What I am suggesting is that we might be willing to take a leap of faith, and come down on the other side of the line, saying that, by implication, this is likely to be validated. That is, all the evidence seems to point in that direction. So right now we are going to label it vocal fold contact in the sure and certain hope that a few years down the road someone will have the evidence to assure us that we were right in doing so. If, in fact, later validation studies show that the validity is poor or absent, the whole technique will disappear from the scene anyhow. So we will be able to bury our error.

Bless: In the meantime, it seems that even just saying that until validation studies are done that no more than a label such as contact can be used might help the reader who does not understand why it has been left unitless.

Baken: Let's be clear: a unitless tracing is the norm in many physiological studies. We do not label units on the electrocardiogram. It just does not matter. I do not know that labeling the EGG trace as square millimeters or whatever means anything given the different sizes of vocal folds. It is the percentage of contact—relative to maximum and minimum—that is likely to be significant. Assigning a unit may not actually achieve anything of clinical utility.

Kotby: So, we would like to stress that the electrical output during EGG registration does not represent glottal area, but it does represent vocal fold contact, which is not to be measured in actual units.

Bless: If we agree it represents vocal fold contact that is not measured in units but think it should have some name, perhaps we could call it contact index.

Laver: From a speech science point of view, I feel uncomfortable with this latest part of our discussion. I think it is important to be specific about what the property is that is being depicted on the vertical axis, and to indicate that it is conductance, in the physical sense. The interpretation is of contact area or other aspects of contact. That is important. Now, if one does call it conductance (which is likely to require some explanation for people who see that term for the first time), that at least concentrates research attention on the fact that there is more to interpretation of that conductance than simply the matter of the contact area of the vocal folds, with all the periglottal aspects of the situation as well. So that, for people to reach a clear interpretation of what they are looking at, and what reliability they can place on the interpretation of that record, it is important that they do not think that there is an immediate relationship between that curve and what they primarily want to know about in terms of the contact of the vocal folds. My own feeling then is that it is better to take a half step back, and call it conductance or some equivalent term, and allow the EGG to be the variation of some aspects of conductance with time.

Titze: I personally have no problems with calling it conductance. But I think that at this point I would probably rather go with contact or contact area, coupled with a strong statement that the validation remains to be done. At the last Tokyo meeting, Scherer and his colleagues reported the results of a study in which vocal fold contact against a sheet of conducting glass was actually measured and compared to the EGG. There seems to have been a highly linear relationship between the two measures, except at the extremes of the function. So I think we can at least guess that, in the future, at least over a significant range of the device, we will see a linear relationship between actual contact area and the output of the device. So I would agree with Dr. Baken's evaluation that maybe we should look toward the future, and expect there to be a calibration for the EGG, and that it would be a legitimate transducer like any other that we buy off the shelf.

Laver: But it is crucial that whatever is said about

contact include the caveats about the need for further calibration.

Titze: I agree that that is crucial.

Bless: Then am I hearing that, with appropriate caveats, we could all accept the term contact? If so, let's accept that term.

Hacki: But there is more to the problem than that. For instance, what should we call the phases of the curve that the EGG provides?

Bless: I believe that question goes a step beyond what I hear being reported here. We may need to deal with it in the future when more data have been collected. Right now, the proposal is not dealing with breaking the curve down into its parts, and deciding how those parts should be labeled, but rather we are simply proposing that one axis be labelled time and the other contact. Is that a fair summary?

Kotby and others: Yes. We should not go beyond that for the time being.

PHONETOGRAPHY

Phonetography defines vocal limits in terms of frequency and intensity. Frequency range is displayed on the horizontal axis and intensity on the vertical axis as shown in Fig. 1. It is purported to be a useful tool in the hands of a well trained phoniatrician, speech pathologist, or laryngologist. It can help identify voice potential in singers; it can be used as a diagnostic aid with dysphonic patients; it can be used to monitor results of treatment; and it can be used to monitor the effectiveness of voice training. Phonetograms can be generated by hand using a pitch pipe and sound level meter, or by using a computer-based instrument designed specifically for creating phonetograms. Patients are asked to produce their loudest and softest voice at each half tone step covering their entire pitch range. Using this procedure it takes approximately 45 min to create a phonetogram that defines a speaker's entire frequency and intensity range. In the interest of time some clinicians have modified the procedure such that they record the maximum and minimum intensity levels at the lowest and highest frequencies the speaker is capable of producing and at four notes around the habitual speaking level. This modified procedure can be completed in as little as 10 min. Phonetography has been recommended as the standard measure of voice by the Union of European Phoniatricians (UEP).

TABLE 1. *Phonetogram data for one patient for 12 different tones, noted by the scale note. This sample represents a reduced protocol from that presented in Table 2.*

			Decibels	
No.	Hz	Tone	Minimum	Maximum
20	147	d	52	52
25	196	g	58	82
27	220	a	60	88
30	262	c1	67	94
34	330	e1	68	93
37	392	g1	74	97
39	440	a1	68	87
42	523	c2	74	88
46	659	e2	87	103
49	784	g2	83	100
51	830	a2	82	98
54	1,047	c3	91	91

Bless: Conclusions that seem to have achieved consensus during our earlier discussion included the following:

1. The phonetogram is a useful tool, whether one is generating it with the help of a computer or doing it with a sound level meter.
2. The instructions to the patient are critical in determining the output, especially in terms of reliability.
3. Comparisons over time and across the duration of treatment in a single patient are likely to be more valid than comparisons across patients.
4. The phonetogram is a useful indicator of the potential for vocal development.
5. The phonetogram must be studied further in order to achieve a classificatory scheme for the different kinds of patterns that might be shown.

Note that the Union of European Phoniatricians has published standards for recording the phonetogram [Schutte HK, Seidner W. Recommendation by the Union of European Phoniatricians (UEP): standardizing voice area measurement/phonetography. *Folia Phoniatr* 1983;35:286–8].

Baken: First, I would strongly urge us to adopt these standards and to recommend them. They do, after all, already exist de facto in the European countries, and it would be well to extend that standardization globally.

Second, there is the issue of the name phonetogram. It is a grossly unfortunate term that is incredibly misleading. It implies, at least in the anglophone countries, that there is some relationship to phonetics or phonetic considerations, which is in no way the case. Phonetogram might already have cur-

TABLE 2. *Phonetogram statistics data obtained for one subject across multiple tokens, illustrating variability in performance*

Token no.	Frequency of production Hz	Corresponding tone	Decibels			
			Minimum		Maximum	
			Average	Deviation	Average	Deviation
17	123	H	58	0	58	0
18	131	c	59	0	61	0
19	139	c#	59	0	65	0
20	147	d	60	0	64	6
21	156	d#	61	1	71	9
22	165	e	61	1	77	13
23	175	f	61	1	79	11
24	185	f#	62	2	82	10
25	196	g	62	3	85	8
26	208	g#	62	3	89	7
27	220	a	62	3	93	6
28	233	a#	63	6	92	8
29	247	h	63	10	91	8
30	262	c1	64	13	90	9
31	277	c#1	64	11	90	10
32	294	d1	64	7	92	11
33	311	d#1	65	5	92	11
34	330	e1	65	2	93	12
35	349	f1	65	4	94	8
36	370	F#1	65	6	95	5
37	392	g1	66	8	96	1
38	415	g#1	64	7	94	6
39	440	a1	63	7	91	10
40	466	a#1	66	10	92	6
41	494	h1	69	11	94	1
42	523	c2	72	14	95	3
43	554	c#2	75	13	96	1
44	587	d2	76	13	96	3
45	622	d#2	79	12	97	5
46	659	e2	80	11	97	7
47	698	f2	84	15	98	8
48	740	f#2	86	18	98	8
49	784	g2	90	22	99	9
50	831	G#2	88	18	95	9
51	880	a2	86	15	90	8
52	932	a#2	85	11	88	6
53	988	h2	83	7	87	2
54	1,047	c3	80	0	85	0
55	1,109	c#3	81	0	85	0
56	1,175	d3	83	0	85	0
57	1,245	d#3	84	0	86	0
58	1,319	e3	86	0	86	0

Area = 701 ± 55 decibels × halftones; dynamic range = 35 ± 0 decibels (averages ± deviations).

rency in some regions, but at least in countries such as the United States, where this measurement is almost unknown among clinicians, we have the opportunity to introduce a better term before this one is widely adopted. I believe strongly that we should do so.

Bless: I thought that we had two different suggestions for names yesterday. One was voice profile and the other was frequency–intensity profile. Stimmfeld, which translates as voice area, is already in common use in German-speaking countries.

Titze: If we are concerned with the boundaries of vocal production, then perhaps the best term (at least in English) is voice range. The problem is that this is likely to be confused with vocal range, which considers only frequency. The phonetogram assesses the frequency-dependent intensity range.

I would like to suggest voice range profile or its equivalent in the language of the term's user. It is

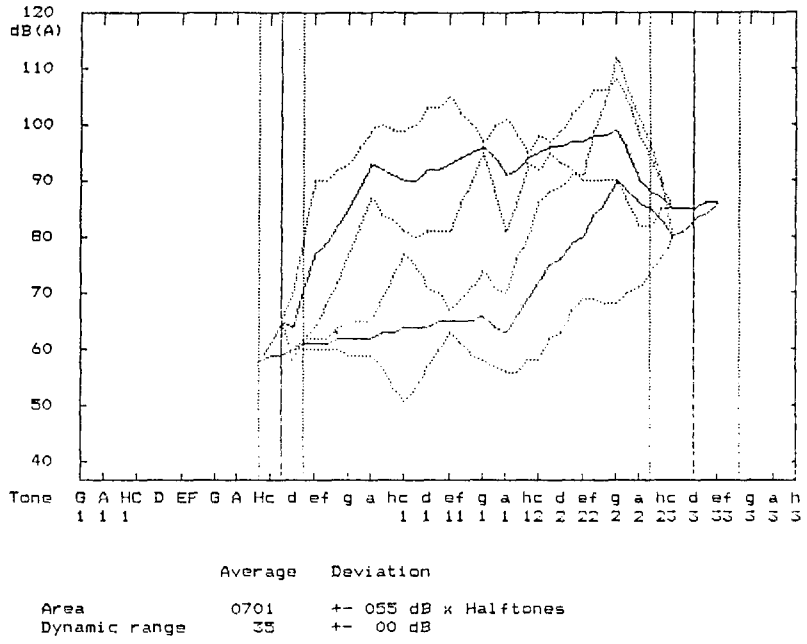

FIG. 1. Diagram of a phonetogram with associated statistical derivations. The heavy line demarks the vocal area before treatment. The light line demarks the vocal area after treatment. The area between the heavy and light lines illustrates the vocal gain.

not as generic as voice profile, a term that is used to label the output of a lot of different kinds of computer analysis of voice.

After considerable discussion, these recommendations were accepted by the group.

ACOUSTIC WAVEFORM PERTURBATIONS AND VOICE DISORDERS

Perturbation is defined as random deviations from complete regularity of the laryngeal waveform. Perturbation of the cycle-to-cycle duration is called jitter. Perturbation of the cycle-to-cycle amplitude of successive laryngeal pulses is called shimmer. Perturbation can occur across a wide range of incidence, slight perturbations occurring occasionally in all speaking voices. Jitter and shimmer contribute to the rough perceptual effect usually called harshness, but perturbations below ~1% are not audible as such.

When patients complain of harsh or hoarse voices, an important clinical decision must be reached. Any objective methods for supporting such a decision, if they can be made automatic, reliable, and cost effective, are of evident value. There are four major applications of automatic methods: (a) screening limited at-risk populations, (b) assessing patients with complaints of harshness or hoarseness, (c) instituting diagnostic support and (d) monitoring to assess changes. There are promising signs that techniques are in principle capable of providing such support, but a major obstacle is the unavailability to date of adequate databases of speakers with disorders and of comparable control groups.

Advantages of acoustic techniques for perturbation analysis are that acoustic recording techniques are noninvasive and use fairly low cost technology; hard-copy records are easy to interpret and can be made easily. Disadvantages are that acoustic recordings are subject to contamination by environmental factors and equipment-derived artifacts. Also, the process of automating the analysis currently involves relatively high cost computing equipment (although it will eventually be simple and cheap). The ideal acoustic recording equipment should have (a) flat, low-frequency response, (b) no signal distortion, (c) high signal-to-noise ratio, and (d) affordability. The electret type of microphone is most useful. Pulse code modulation (PCM) digital audio recorders provide a flat-frequency response over a wide-frequency range, with minimal phase distortion. Several relatively inexpensive analog recorders offer flat frequency response. However, analog magnetic tape recorders produce phase distortion caused by amplifiers and preamplifiers. The issue of recorder type has recently been circumvented by direct analog-to-digital conversion of speech signals onto digital computer systems. Re-

cordings preferably should be made in a sound-treated booth which also provides shielding against main current electrical signals.

Three types of databases are needed: (a) voice samples from specific voice disorders; (b) speakers recorded before diagnosis with information about the later diagnoses; (c) a control group divided into smokers and nonsmokers. Neither should have a history of voice disorders. The voice samples in the databases should include not only sustained vowels, but continuous speech of ~40-seconds duration.

The major issue in perturbation remains the correlation between the acoustic characteristics of perturbed phonation and the physiological and aerodynamic situation to which they correspond. So far the search for an automatic system providing diagnostic support to clinicians is limited to correlational studies of the connection between types and degrees of perturbation and the typology of disorder that laryngoscopic inspection and histology can show.

Finally, there are many ways in which one can mathematically treat the variation about an Fo trend. Even though various formulae have been presented in the literature, there is no consensus at all about how perturbation should be mathematically defined. Some will be sensitive to low-frequency variations, others will be sensitive to high-frequency fluctuations. Still others will demonstrate overall average change. The measures must be tuned to the specific disorder that causes the perturbation. Until we determine the mechanism of how a certain perturbation is produced, we cannot fine-tune our measures. We therefore are appreciative of the fact that Dr. Laver left open the specific definition of perturbation. Those who write journal articles will have the obligation of clearly defining the specific measure that they have chosen.

Laver: There is obviously a great deal of work to be done in the acoustic contribution to phoniatrics, speech pathology, and laryngology. But I think the single most important statement we can make, springing from the perturbation work (just as an example), is that acoustic data are of relevance to phoniatrics and laryngology. That is a message that may have been received by the converted in this room, but it is a message that has clearly not yet been accepted by the profession at large. So that the usefulness of acoustic data as support data, and as quantification, as well as a means of adding qualitative information to the armamentarium of the profession is extremely valuable and must be stressed strongly. So that, for example, the training curricula of specialist in these areas need to address the problem of how to understand and interpret acoustic records so as to gain maximal advantage from the extra information they provide. The value of acoustics to the profession must be emphasized to a greater degree.

Baken: Might I ask that we insert into this discussion of perturbation some notice of the fact that there are also nonrandom perturbations? So far, unless I have misread the discussion badly, the assumption has been of essential randomness or quasirandomness. But there are patterned deviations, particularly resulting from tremor which is the most obvious example. The methods that evaluate perturbation in the random sense can, with less than enormous changes, be used to detect the magnitude and rate of tremor. This can be invaluable in clinical practice.

Laver: There is a fuzzy semantic boundary to the concept of perturbation. To most people, perturbation has a necessary random element. I think that is not helpful in this area. What we need to do is look at the signal characteristics of successive pulses. They can form all sorts of relationships, some of which seem to have a quasirandom basis, but others of which are more orderly but will nevertheless show up in some the equations used to capture perturbation. So ould recommend being all inclusive rather than trying to reach a narrow definition of perturbation that says that vibrato, for instance, is not an aspect of this phenomenon.

Titze: Dr. Laver, your paper is very interesting, and raises several questions. First, I am wondering how we separate prosodic trends from low frequency tremors in, for example, neurologic disorders? I am somewhat surprised to hear that jitter values below 1% are normal. That seems like a high threshold in light of recent findings, that normal jitter can be as low as 0.1%. Much of what has been previously reported as vocal jitter may have been equipment jitter. My third question concerns electret microphones. Are they really the best for perturbation analysis? They seem to be susceptible to low-frequency noises. Perhaps some dynamic microphones with response flat above 40 or 50 Hz would filter out some of these low-frequency rumbles.

One of the problems of standardizing recording

equipment is that the industry is in a state of revolution. Nobody seems to know what type of tape or disk medium will prevail. PCM audio processors, although ideal for our purposes, are hard to find. We must wait for some stability in the commercial market.

Additional discussion concerned how clinicians can make high-quality recordings inexpensively. The committee recommended that a subcommittee be formed to outline directions for standardizing recording procedures. The committee agreed that it was premature to specify the type of acoustic analysis to be accomplished because at this point in time it is unknown which acoustic parameters are most relevant. One suggestion was to encourage development of analysis techniques that could make use of telephone lines to screen for pathology.

INTEGRATED AERODYNAMICS

Loebell: Would you elaborate on the relation between incomplete glottis closure and subglottal pressure?

Schutte: In incomplete posterior glottic closure we found either higher Psub or "normal" pressure sometimes accompanied by higher than normal airflow rates. The aerodynamic data in these patients resemble comparable data in patients with vocal nodules, which also show an incomplete dorsal glottis closure. After voice training the subglottal pressure decreases but the airflow stays more or less the same. We interpret this to mean that subglottal pressure measures are useful in monitoring treatment effects in selected pathologies.

Sonninen: How do you explain the occurrence of high subglottal pressures with incomplete glottal closure?

Schutte: It appears to be the result of a type of counterbalance between laryngeal musculature and the breathing apparatus that supplies the air pressure. The counterbalance is set at a high level, either by a too high laryngeal muscle tension or by a too high subglottal pressure caused by excessively high activity in the breathing mechanism—no breath support. In other words, there may be a close relation between high subglottic pressure and high laryngeal tension.

Kotby: Can you relate this to the pathogenesis of the disease? Do you think the subglottic pressures is a cause or result of the glottal waste?

Schutte: That is a difficult question to answer because the two are so interrelated. You could easily make a case for either side.

Baken: I would like to comment on two things. First, the larynx and the chest wall are independent; the pressure is generated by the chest wall system, which in essence can generate whatever pressure it wants. In theory, the system is regulated in the engineering sense so that it should be able to maintain a pressure within reasonable limits despite the magnitude of the airflow. Although it is not ideally regulated, it should be able to maintain a pressure with the cost being most apparent in a reduced amount of time for which one can keep an airflow going with a high pressure and leaking glottis. The pressure should be relatively independent of the status of the glottis if the time factors are ignored. This brings up the second issue I wanted to raise concerning subglottal pressure measurements and diagnosis. The pathology of interest is not only in the larynx, but should include anything within the organism that is likely to change voice production. For example, the neuromuscular disorders can effect motor stability; the basal ganglion disorders in particular can have a dramatic effect on voice production. Thus, an individual's ability to maintain a stable subglottal pressure during the course of a speech utterance may be much more informative than the actual pressures generated. Thus, I would like to put in a pitch for looking at subglottal pressure, not necessarily in terms of the pressure adequate for phonation—because I agree that people use a wide range of subglottal pressures—but rather in terms of the time course of stability.

Schutte: I agree, but still have the question of how to decide which of the many ways of measuring should be used in this type of patient. For example, measuring the mean pressure may not give the proper information, some of the phenomena occur at a specific frequency, perhaps 10 Hz. You must be able to track fast changes, which presents a problem when using the esophageal balloon. Using a minipressure transducer between the vocal folds probably cannot be used with many of the neurologically involved populations, leaving us with subglottal punctures.

Baken: I am not totally in agreement. If one is not looking for an absolute calibration of the subglottal pressure (that is, when one is not interested in its exact value to the nearest tenth centimeter), then a wider number of options are open for estimating subglottal pressure. When the primary interest is stability, or change in subglottal pressure over time,

then several techniques might be useful (e.g., either the intraluminal technique using a miniature pressure transducer or body plethysmograph). Obviously, it is open to debate, but my sense is that the change over time is the variable of interest.

Wendler: Schutte made a very important distinction in his paper between aerodynamic measures used for voice physiology research and those used for diagnostic purposes. Many of the measures, including aerodynamics, are useful for the severely involved patient. Is it possible to use aerodynamic measures to differentiate in subtle voice changes? I believe we should underscore a passage from Schutte's paper for this conference, ". . . because of its weakness for diagnostic purposes aerodynamic measurements are unlikely to be useful in routine daily clinical practices."

Hirano: The final goal of application is clinical use, but for the future we must do additional research to determine its best use and need for further development.

Wendler: Is use of the esophageal balloon clinically practical? Do patients actually accept the balloon?

Schutte: They do; it offers no real problems. "Experienced" patients swallow the balloon even without anesthesia. A fairly flexible catheter is inserted through the nose until the clinician feels a small resistance at the entrance of the esophagus.

Titze: Martin Rothenberg apparently uses some kind of compromise by having his subjects swallow a catheter with an attached miniature transducer. Could you comment on the advantages and disadvantages of this technique in terms of patient comfort, ease of administering, and data interpretation?

Schutte: Rothenberg uses a miniature pressure transducer within a water-filled catheter placed in the upper part of the esophagus. The data, measured at that point, are difficult to interpret because the position of the transducer is critical and somewhat difficult to ascertain. Moreover, to obtain data comparable in frequency band with data obtained directly in the subglottal space, the pressure transducer in the water-catheter should be positioned very high in the esophagus, where a higher muscular tension normally can be found, the upper high pressure zone in the esophagus. This may interfere with the measurements. So, in answer to one part of your question, the esophageal placement is preferable for data interpretation. As far as the patient comfort is concerned, anything placed between the vocal folds, compared to within the esophagus because of the cough reflex, is likely to be more of a problem.

Titze: Let me ask one follow-up question. If you were going to do an experiment, a control of intensity, and you were to combine subglottal pressure measurements together with electromyography (EMG) of at least two or three muscles, with normal subjects and singers, which method would you use?

Schutte: If accuracy is the major concern, the transglottal pressure measurements would probably be preferable. Of course, a potential problem would be interference between the electrodes and the catheter.

Hirano: I agree with others who have suggested that aerodynamic measures are not very useful for determining the cause of the disease of the voice disorder. Skilled clinicians have the ability to make a diagnosis by integrating information obtained with their eye, ear, finger, fiberscope, stroboscope, and in some instances EMG. Nevertheless, in Kurume we routinely use aerodynamic measures; our purpose is to measure the degree of dysfunction. Remember that no single test can measure total vocal function, we must test several different aspects to get a composite picture of how the laryngeal and respiratory system are working together to produce voice. Thus, aerodynamics provides information on one part of the total picture. Another use that should be mentioned is in patients with glottal incompetence where treatment can be monitored by airflow measurements.

Baken: Apparently there are some international differences in the use of the word "diagnostics." In American it is used to describe what is the cause of the aberrant voice in terms of what in the mechanism has failed, and in what way, irrespective of a specific diagnosis. Confronted with a grossly abnormal voice, it is important to understand what parameter of vocal function has been disordered. For those purposes, for those diagnostic purposes, aerostatics and aerodynamics may be relevant.

The committee agreed to adopt the broader meaning of diagnosis to include the assessment of the degree of functional disorder.

Hirano: Using the broader definition, aerodynamic measurements are useful for diagnostic purposes.

The committee agreed.

Loebell: Would you agree that the direct subglottal puncture is still the best way to obtain subglottal pressure measurements? The technique is simple, you do not have to be so concerned about the location of the transducer, and it is relatively easy on the patient. Aside from a large thyroid gland I do not know of a contraindication.

Schutte: I would fully agree, and believe it is also suitable for research purposes.

Wendler: Aerodynamic measures do not provide much information on etiology, but aerodynamics are useful in classifying the degree of functional pathology, that is, to separate organic from functional.

Hirano: For a given glottal configuration there is an optimal pressure; the pressure can be too high or too low to make easy vibration. With an incompetent glottis and a high pressure, the flow will also be higher but there are many different situations that can result in similar aerodynamic patterns. For example, too much resistance can be caused by too much muscle contraction or by thick material such as scar tissue.

Fex: I have several questions. Is there general agreement that aerodynamics are useful for evaluation of treatment effectiveness? Are pressure changes active or passive, controlled by pressure receptors or some other mechanism?

Schutte: For clinical purposes, mean pressure is generally used. There are both active and passive components. The issue is relevant to discuss because the myoelastic aerodynamic theory of voice production states that the glottis opens when the pressure becomes sufficiently high to overcome the resistance offered by the vocal folds, the closed glottis allows the rebuilding of the pressure, and so on. In our aerodynamic and electroglottogram (EGG) measurements of singers and nonsingers, we found that at the moment of glottic closure the pressure peak is highest. Thus, the peak coincides with the beginning of closure. We also noted a difference between trained and untrained singers. The best singer benefits from this high pressure point.

Baken: Rapid DC changes in vibration cannot be controlled by mechanoreceptors because they are too slow to follow the fundamental frequency of the voice.

Titze: There is little doubt that the subglottal pressure peak comes at the point of glottal closure. In fact, glottal closure produces this peak. As the vocal folds collide, the air column momentum in the trachea compresses the air below the glottis and excites a large acoustic pressure. Supraglottally, a similar air column momentum produces a negative pressure. The result is a large transglottal pressure before glottal closure. Contemporary theories of vocal fold vibration do not depend on negative Bernoulli pressures before closing, as formerly thought. (For a description of the mechanism of vocal fold vibration, see Titze's discussion in *JASA* February, 1988).

Sonninen: Why is there a drop in pressure during closure?

Titze: Psub and Psup are out of phase.

Loebell: Pressure and flow measurements should not be limited to means during phonation. From a clinical perspective, voice attacks/voice onsets are equally important.

VOCAL EFFICIENCY

Problems

The concept of efficiency is grounded in the field of mechanics. In that domain, its definition is relatively simple and its utility clear. Its applicability to voice production is somewhat more clouded, however, and issues of vocal efficiency are less easily dealt with.

Vocal efficiency is not synonymous with laryngeal efficiency: turning of the vocal tract plays an important role in determining the radiated acoustic power. Acoustic loading on any source can improve its efficiency, and it is reasonable to assume that vocal tract characteristics can be adjusted to optimize this effect. Ideally, efficiency measures should take into account power losses in the laryngeal musculature (for example, in antagonistic contractions) and in the chestwall system. A major problem, then, is obtaining estimates of the components of the overall efficiency.

Efficiency, in any case, is not the same as vocal effectivity, which may be more important from a clinical perspective. Yet this is a parameter difficult to define and perhaps impossible to quantify.

Finally, and perhaps of paramount importance to the issue of clinical application, is the fact that measurements of efficiency do not speak to issues of the long-term health of the vocal system. Short-term gains in energy conversion might easily be obtained at the price of eventual injury or disorder. Thus, great caution is advisable.

Recommendations

1. Because oscillator efficiency is directly proportional to the square of the oscillation frequency,

measurements must be taken at several rationally selected and reproducible relative frequency levels.

2. Within a restricted range, efficiency also tends to increase with intensity. Hence, standardization of test intensity levels is also important.

3. Because efficiency might change in meaningful ways as a function of speech–task duration, it will be useful to develop test procedures that are the vocal equivalent of treadmill tests, with multiple determinations of efficiency taken as the test procedes.

Needs for technologic development

Aerodynamic power measures remain difficult to obtain. Estimates of subglottal pressure and flow based on consonant-vowel-consonant (CVC) utterances (/pæp/) or airway interruption techniques are acceptable when carefully used, but better methods will be needed if efficiency testing is to become more valid and convenient. It would be very helpful to have a method for evaluating the translaryngeal pressure loss, which would allow determination of glottal efficiency alone.

Data-base needs

Clearly, much more in the way of normative data are needed, and a more solid understanding of the efficiency losses associated with laryngeal disorders is required.

Titze: I agree very much with the summary. I would just add that, with respect to the issue of holding frequency and intensity constant, we might not have to be so rigid as to say that it has to be the same intensity and frequency for all subjects. But for a given subject, I would say that if repeated measures of vocal efficiency are to be made pre- and postintervention, then they should be made at the same frequency and intensity level. And one way to do that is perhaps to choose a strategic point in the phonetogram (frequency-intensity range profile) made on the patient at the time of initial testing. Such a point might be right in the middle of the phonetogram, at the point of crossover of the major and minor axes. One could decide that at this point, efficiency measures will be made. The same point could then be used for follow-up efficiency measures.

Hirano: I certainly agree with Dr. Titze's proposal concerning the selection of a frequency and intensity standard for repeated testing. But in clinical tests, it sometimes happens that the pitch and intensity ranges change in the course of treatment. For example, in Reinke's edema, posttreatment, the pitch will rise. In cancer patients treated surgically, pitch range may be much reduced. In these patients we must have some other alternative for choosing the frequency and intensity at which efficiency measurement is to be done.

Titze: But in those cases we should expect that the vocal efficiency values will change, is that correct? So maybe the whole measure at that point might not be that meaningful.

Hirano: In the case in which most comfortable pitch and loudness change after treatment, it might be most valid to consider the efficiency at the new comfortable levels for comparison with the efficiency at the old comfortable levels.

Titze: Yes, I understand what you are saying. I guess what I am thinking then is that if the levels of comfortable pitch and loudness change before and after some kind of intervention we should make a note stating that vocal efficiency should be totally recalibrated, and that we can not make a very good comparison before and after treatment.

Hirano: I am not sure whether we are justified in saying that we cannot make a good comparison.

Titze: Well, we can write numbers down. But I am not sure what they mean.

Hirano: Well, the most comfortable pitch and loudness may be associated with the least muscular energy consumption.

Baken: I do not know that this is really a problem. Let's assume that you establish a frequency–intensity vocal profile (FIRP), say in your edema case. Suppose that the standard for testing that we ultimately evolve requires, just for the sake of example, three frequency levels at a constant intensity. For the moment, let's call them low, medium, and high. Then you test your preoperative patient at these frequencies and a constant intensity (or perhaps even at several intensities), you do your intervention, and a month later you retest using the same frequencies for low, medium, and high. Now, what we would expect to happen is that there would perhaps be a diminution of efficiency at the low frequency, because the vocal range has been shifted upward, and increased efficiency at the medium frequency, and more efficiency at the high frequency. Or some pattern like that. But if you had originally sampled over the individual's range, even though that range had moved, on retesting at the same frequencies you would get an image of what had happened to the patient's efficiency across his range. So it might not matter that a patient's range had shifted.

Titze: Yes. If you measure at a sufficient number of points. You are suggesting three and, hopefully, within those three points you would see a trend in that efficiency would fall at some point and rise at another point. We have not made any recommendation about how many points, say within an FIRP, one should test at, but if you test at a large enough number of points then I think what you are saying makes a lot of sense.

Bless: So it seems as if one of the things that will obviously need to be a consideration in the future will be working with glottal efficiency relative to an FIRP, and integrating those two kinds of measures. Speaking from a strictly theoretical point of view, I think that Dr. Baken is quite right. But in the practical realm of the clinical setting, patients must submit to an awful lot of testing. And there is a limit to the tolerance of the patient and the third-party providers. If the patient can tolerate all of the other required tests, testing efficiency at multiple points in the patient's range is ideal. But if not, there is a problem. This sort of thing is very important in clinical settings. We cannot spend 10 h testing a single patient. And the testing may be useless, from the patient's point of view.

Baken: My image of all this testing may be different, and if it is not a valid picture, then I hope you will tell me. My picture is not that every test that could be devised will be applied to every patient, any more than every conceivable blood study is applied to every patient who is admitted to a hospital. Rather, based on clinical judgment and insight, those examination procedures that are most likely to show the nature and severity of the problem would be used with a given patient. If in fact it seems reasonable to the clinician that an efficiency study would provide a great deal of information in a given patient, that test assumes a higher priority and we use up the patient's tolerance for testing in descending order of test priority. I am less concerned that we may generate recommendations for dozens of tests because I would never expect us to say that all patients should be run through a battery of this exhausting and boring kind of material.

Bless: Shouldn't that caveat be stated up front, so that clinicians clearly understand that that is the case? Should we not also be considering recommendations for a basic battery of tests to be administered to every patient, perhaps as an initial screening, to serve as the bases for making clinical decisions about branching for the additional procedures that are likely to produce the greatest clinical yield.

Kotby: From a practical point of view, I think we will find that most of our mysterious functional voice disorders will have to go through most of these tests because we simply do not know much about the problem and sometimes must make the diagnoses by ruling out other problems such as myopathies and neuropathies. Is ruling out sufficient justification to subject our patients to all of this testing? Another related question concerning what our patients should be subjected to is whether we should still perform direct measurement of subglottal pressure by tracheal puncture. From this committee there seems to be some positive support for use of this procedure.

Titze: I think we concluded that unless you really do make power measurements, and really refer output energy levels to the input, then even labeling it efficiency is not correct. AC/DC ratios may give you something about collateral competency, but they certainly should not be called efficiency measures. By the same token, I think if making the subglottic pressure measurement entails a real overload in terms of time and energy then given the fact that the whole idea of glottal efficiency is only a very small portion of the laryngeal efficiency or vocal efficiency, I think it is a judgment call as to whether you want to go through with the procedure.

Baken: I heard the question somewhat differently. I think we might add a notice of tracheal puncture or whatever. But it was my understanding that we are discussing needs for further development. Methods for tracheal puncture do exist and are available to any clinician who chooses to use them at the moment.

Kotby: But is the data that we extract from direct equations in which the subglottic pressure was very important reflecting much that is of clinical value.

Titze: I think they are very preliminary if they do have clinical value.

Hirano: Perhaps things have gotten a little complicated. We need to make them simple. Basically, there is no objection to measuring efficiency at three or four frequencies on a patient. In those cases where frequency levels do not change before and after treatment, a single measure can be used. You have to try to keep the pitch and intensity constant in order to make a clinical comparison. But in those particular cases in which pitch and intensity are markedly changed, as a result of treatment, as a second choice you can choose a specific way of selecting test levels such as a comfortable pitch and intensity at piano, forte, and mezzo-forte.

Titze: What we can add is that if the clinician feels that the overall range of intensity and frequency have dramatically changed as the result of intervention, at that point efficiency should be measured at more than one location in that range in order to get a meaningful reassessment of glottal efficiency. Otherwise, perhaps a single measurement at a standard frequency and intensity would suffice.

Hirano: I still believe that it is useful to measure at a single frequency and intensity level if that level is the most comfortable for the patient.

Bless: Couldn't we say that efficiency testing would always be done at the most comfortable level, whatever it is, but if the pre- and posttreatment sessions differ by more than a given amount, which we would have to specify, that multiple measures should be used?

Titze: How will we determine what change in efficiency was due to intervention and what was changed by some more general change in overall conditions that occur in the voice. I am a little bit lost as to what we are trying to determine.

Hirano: We are not dealing with machines, but with human beings, so we cannot always determine what caused the change in efficiency. The challenge is to document that a change has occurred.

Baken: The problem seems to be that we are all examining an animal that has not been fully unveiled yet. It is clear that we need some clinical experience in the application of efficiency measurements. We also need improvement in the measurement techniques. And until there is some real experience in the real world, and until we come to apply this, it seems to be premature to begin talking about specific frequency levels and how reproducible repeated measures are, and all of that. So far as I know, with the possible exception of Dr. Schutte's experience, I do not believe anybody has heavily applied this to a clinical population. I think one of the reasons that we are having such trouble in deriving recommendations is that we do not have the experience and the base of patients yet to begin making these decisions. Considerably more work is going to have to be done before we can specify recommendations for this procedure.

Sonninen: I have the impression that in clinical work we have different kinds of programs. One takes a wider approach, another a more narrow one. Only in limited cases is testing done very deeply. Have you some experience or insight about which cases one would recommend these efficiency measurements? For every patient or only for those who present with special kinds of problems?

Hirano: That is related to what Dr. Baken said. No one has an exact answer at this stage. Many clinicians empirically know what test should be the first priority for a particular case. For the time being we have to leave it to the clinician's judgment because of the diverse availability of resources to different clinicians and clinics. But once you decide to do a particular test, for example, we would hope clinicians would follow these recommendations that we have been working on.

Baken: Can I try to rephrase what I did not say well before. My view of our work now is to generate a report that says that the possibilities for examining vocal function are rapidly broadening. That technology is making more things possible for us. At this stage, and after this much deliberation, our report should say, "This is our collective wisdom about what is likely to be beneficial, about what we feel is probably valid, about what types of testing are feasible in the clinical environment. While, for most of those tests, we have not enough clinical experience nor is theory necessarily well enough developed for direct application today, and while we cannot yet write recommendations that are specific in many areas, we are sort of alerting the professional community: These are issues worthy of further development and attention." That, I believe, is as far as we need to go in our report.

Fex: In short, we need more basic knowledge. And when that knowledge becomes available, then things could be specified to be done in a particular way.

Baken: Or, these are the considerations that the clinician should keep in mind.

Kotby: As a practical solution for the moment: If I have a case for which I expect that the vocal pitch will ultimately end up an octave lower or higher after my intervention, then why not have the preintervention test at two pitches: the present and the anticipated? Perhaps, if I expect pitch to fall by an octave, I should test at the present comfortable pitch and at a frequency one octave lower.

Titze: Might that not be too much guesswork?

Kotby: But we expect that certain conditions will improve. Because if the pitch is highly reflective of the pathology, then previous experience serves as a basis for an informed judgment of the amount of change we can expect. So I would prefer to test at frequencies that anticipate the change.

Bless: It would be an interesting task for the com-

mittee in the future to go through some decision making for each of the types of testing procedures that are done, given a certain population of disorders. We might consider what would be the decision making that would go into using or not using a test and the anticipated results. But for today, that task seems premature, especially because we do not have the database upon which to proceed.

Hirano: Another possible solution to the future, based on Dr. Schutte's figures, would be to determine if there were some way to normalize the pitch and intensity. This would be very important in cases in which the pitch level changes significantly.

Titze: Yes, that is quite possible. One might have a nomogram developed by which one could readjust all the values.

Loebell: May I remind you of one sentence of Dr. Schutte's presentation: "Because of their weakness for diagnostic purposes, it hardly can be stated that aerodynamic measurements will be useful in daily routine practice."

STROBOSCOPY

Stroboscopy is a highly relevant qualitative method for observing vocal fold movements. It has proven to be an essential component of examination of the larynx. Stroboscopy is particularly valuable for the early detection of malignant pathologies (nonvibrating portions of the vocal folds), identification of vocal fold cysts, evaluation of treatment effectiveness, detection of early recurrence of malignant lesions, identification of movement problems resulting from laryngeal trauma, description of vibratory patterns resulting from excessive or inadequate muscle function, verification of return of function in cases of laryngeal paralysis, and for provision of a laryngeal image for visual feedback therapy. Consequently, laryngostroboscopy should be an obligatory part of systematic outpatient care programs for cases of chronic laryngitis. Additionally, stroboscopic investigations are strongly recommended before, after, and, when possible, during phonosurgery procedures (indirect microsurgery under functional control). Stroboscopic recordings should also be considered a necessary prerequisite for expert opinions provided for legal purposes. For the majority of cases in routine clinical practice, the rigid endoscope is recommended. Flexible fiberscopes provide a more comprehensive image of changes occurring during speech and singing but have poorer optical resolution. Amplitudes, mucosal waves, and phase of closure are the most relevant parameters to observe. Irregularities can be observed but cannot be measured precisely without coupling the stroboscope with additional instrumentation. To replicate and interpret the laryngeal image require that pitch, intensity, and mode of phonation be controlled. Frequency and intensity should be noted for all recordings. To obtain a comprehensive impression of the dynamic variability of the vibration pattern, stroboscopic investigations should be systematically conducted at different vocal intensities. For most purposes subjective quantification of the laryngeal image is clinically satisfactory. Objective, absolute measures are available but are by no means necessary, and with current technology are frequently both tedious and time consuming to complete. An underused stroboscopic technique is microstroboscopy, which provides high magnification and an excellent stereoscopic view. Clinicians are urged to use this technique, particularly in cases where standard stroboscopic examination does not show any cordal dysfunction. Manufacturers are implored to continue to develop their instruments: to include both frequency and intensity outputs; to increase the light intensity; to couple with inexpensive acoustic, electroglottographic, and aerodynamic systems; to increase the frequency response; and to improve the filters.

Bless: I would like to add that international differences in scope of training and practice may, in part, dictate what we can specify. Because things are done differently in various countries, I believe it important to underscore your suggestion that stroboscopy should be done by whomever is responsible for evaluation. So, for example, in the United States we do not have phoniatricians; consequently, stroboscopy might be done by a speech pathologist, a speech scientist, or a laryngologist.

Loebell: We need to temper that a little. In Western Germany, for instance, there are speech pathologists and speech teachers. It would not be appropriate for speech teachers to undertake stroboscopic examination because they have no training at all in these things. So I agree we should be very careful with respect to international issues of that kind.

Bless: So it might be best not to specify, but just to state that it must be done without saying who should do it. As an alternative to specifying specific professions, we could provide a list of minimal qualifications.

Kotby: I would like to add something relevant to oncology. Since its development, stroboscopy has been the "magic tool" for early detection of tumors

of the vocal folds. When we have tried to convince laryngologists to use it, the answer we would often get was that they see their patients when the tumor is too advanced to use stroboscopy effectively. We might agree with them: most cases come to the clinician after the tumor has already acquired significant mass. Thus, it is not useful to discuss early detection. But after intervention, whether it is surgery, behavioral, or laser, the clinician has no excuse for delaying early detection of recurrence. And here stroboscopy might be of great value, not only in detection, but also the discovery of recurrence. We have been stressing early detection of the primary tumor too much. This is less beneficial than early detection of recurrence. After initial treatment, the patient is coming regularly for ongoing observation, and it would be the clinician's responsibility to perform stroboscopy every time to detect any symptom of recurrence.

Wendler: I have an objection to the use of a recommendation in the Faure and Muller presentation. They said that they would like to eliminate the terms hyper- and hypofunctional dysphonia and instead refer to organic findings like submucosal cysts. But submucosal cysts are rare.

Hirano: Cysts are not rare.

Wendler: I find them to be rare.

Hirano: There is a very important point here. I have noticed that in each country the specialists who deal with voice differ. In Japan otolaryngologists are the only professionals who can deal with voice disorders: we have very few speech pathologists, and their training in voice is minimal, consisting primarily of postgraduate workshops. In fact, in all of Japan we probably have less than 25 Ph.D. level speech pathologists, and most of them have specialized in neurogenic speech disorders. But in other countries, for example in European countries, both phoniatricians and otolaryngologists specialize in voice. In the United States they have otolaryngologists, speech pathologists, and speech scientists specializing in voice disorders. So the patient population may be much different for one particular profession in one country than in another. And in my experience, a vocal fold cyst is not a rare condition.

Loebell: May I add that it is still important to provide information about the functional nature of voice disorders toward hyperfunction or hypofunction, because it influences the organization of the treatment. Thus, for treatment recommendations it is crucial to document what is the vocal fold closure pattern of the patient. I consider this to be providing a description of the nature of vocal function, which differentiates hyper- and hypofunctional patterns.

Bless: I think a major point of Drs. Faure and Muller's paper is that sometimes in the past, in the absence of stroboscopy, disorders have been misdiagnosed as functional when in fact there has been an organic basis for the problem.

Kotby: I think that we should highlight the value of stroboscopy as showing not only the pattern of vibration, but also as giving us insight into the morphological pattern of the vocal fold in a way that can never be achieved without stroboscopy. And in this light we are justified in making a recommendation that we do not jump to a conclusion of functional disorder before we use stroboscopy to detect the minor evidence of an organic problem. I believe that this is implicit in what Dr. Bless was saying.

Bless: That is a very good caveat.

PERCEPTUAL ASSESSMENT

During a somewhat lengthy and redundant discussion Fex agreed with the point that the term used for listeners' judgments should not be psychoacoustic assessment, but rather suggested the term perceptual assessment instead. Hirano suggested that unadorned perceptual might not be accurate enough. The compound term auditory perceptual judgment might be a better descriptor. Laver favored perceptual assessment as a perfectly adequate and well-understood term, and Baken added that the sensory channel through which a judgment was made would almost always be clear from context.

Baken: My suspicion is that, in each of our respective languages, in writing up either research or clinical reports, every writer is going to use whatever term is most comfortable. I can tell you for a fact, for instance, that I am not going to change my preferred vocabulary as a result of any recommendation this committee might make. I do not think anyone else will either. What I think we are trying to say is that one should avoid use of the term psychoacoustic, which has other meanings that are much more specifically defined. I think that really is the only recommendation that we can make that has strong merit. We should simply alert people to the fact that there is a group of scientists who are called psychoacousticians. Their field has established certain very rigorous procedures that, in their sum total, are psychoacoustic measurement. We are urging people, when simply listening and forming a rel-

atively crude judgment, to avoid using that term, lest it imply more than what really went into the assessment. For the rest of it, I would be perfectly happy to leave the specific term to the judgment of the author.

Bless: I agree with most of what you have said, except that I think that, to the extent that we can agree and make a recommendation, there may be some people who are less wedded to their own preferred vocabulary. They might be sufficiently flexible to make minor changes in order to make it easier for others (especially less experienced people) reading the literature. If we could come up with a terminologic recommendation, I would urge us to do so.

Committee agreed: The terms perceptual evaluation or perceptual assessment should be recommended for standard use.

ELECTROMYOGRAPHY AND NEUROGRAPHY

Electromyography (EMG) is a useful clinical tool. It provides information about the integrity of the neuromuscular system, and can discriminate between superior and recurrent laryngeal nerve problems, and functional and organic disease. The cricothyroid and thyroarytenoid muscles can be easily accessed with percutaneous laryngeal EMG. The postcricoarytenoid and other intrinsic muscles may require laryngeal telescopic application of the electrodes.

The type of electrode and apparatus selected for EMG studies will be determined by the objectives of the study. Hooked-wire electrodes are best suited for kinesiology studies because there are fewer problems with displacement and less interference with phonation/articulation maneuvers; concentric needle electrodes, with a fixed distance between poles, are well suited for the clinic because of the ease of use.

Use of laryngeal EMG is not routine in most clinics, even when it is indicated. This may be due, at least in part, to problems that have interfered with reliable interpretation: electrode misplacement, electrode displacement, motor unit action potential analysis, and test protocol.

Even less routine is the use of neurography and reflex myography, although these techniques are believed to answer important clinical questions: Where is the site of the nerve lesion? What is the type of nerve lesion? When is the optimal time for intervention? Problems encountered in data interpretation may have inhibited its widespread use as a clinical diagnostic and prognostic procedure.

Sonninen: How do you differentiate whether an EMG activity is real? For example, if we monitor EMG recordings from the tongue while a person is writing we find muscle activity (i.e., pseudokinesiology). In this instance, how would EMG tell you if the muscle activity is really necessary for the function—writing?

Kotby: This is not easy, and one must be cautious in making interpretations. We must take into account in our interpretations potential recording artifacts, the recognition that muscles can function as either agonists or antagonists, and the recognition that everyone does not use the muscles in the same way, or in the most economical manner.

Wendler: How useful is EMG in functional disorders? In other words, how much is the system disturbed by the EMG measurements?

Kotby: Not very much, no more than other methods. Data on functional disorders is meager.

Hirano: I saw a psychogenic aphonia with normal TA and CT muscle activity. She also had activity of the PCA during phonation. This demonstrates the need to sample more than the TA and CT in some populations.

Fex: In your opinion, what is the primary motivation for EMG?

Kotby: It helps to exclude problems.

Wendler: I have two questions: (a) how reliable is EMG in prognosis, and (b) do you use EMG selectively for particular laryngeal pathologies, or do you use it routinely for all cases? We have not found it to be very reliable for predicting the course of paralysis.

Kotby: First, I will address your question concerning prognosis. Second, I will respond to your question concerning the indications for EMG. EMG is indicated in mobility disorders of the vocal fold when we are uncertain of the condition. EMG provides some indication of whether the neural muscle system of the larynx is intact. It provides a means to map the status of different muscles. If there is any indication of regeneration, we wait before treating the patient.

Hirano: We use EMG routinely in our diagnosis of immobile vocal folds. In most instances it is not necessary to do more than the TA and CA because one is innervated by the RLN and the other by the SLN. EMG is the only procedure with which we can determine if the SLN is intact: the position of

the vocal fold is not a reliable indicator. EMG has limited use as a prognostic indicator; if voluntary activity is present in the immobile fold, during the first month after onset the paralysis is partial and consequently the prognosis is more favorable; absence of voluntary activity within 1 month postonset cannot be interpreted; 3–6 months postonset, voluntary action potential is appreciated with favorable prognosis, although not necessarily complete recovery. If a nerve is completely sectioned we see fibrillations, and now we can go ahead and treat the patient rather than waiting a year to see if there is recovery. EMG is also useful in differentiating mechanical ankylosis of the joint from paralysis, and psychological dysphonias from those of paralysis. Fibrillations have a bad prognosis; in that case there is no need to wait 1 year, voice therapy is unlikely to be profitable, and surgery is indicated. Again, we need to be cautious because EMG, although useful, is not 100% reliable.

EMG may also be indicated in trauma cases. I once had a patient who was a player of Japanese fencing. His voice was not hoarse but he could not shout or produce a high pitched voice. His vocal folds moved well for adduction and abduction but he was unable to produce a normal vocal range. There appeared to be two possible etiologies of his problem; he could have a sectioned superior laryngeal nerve or mechanical damage caused by a laryngeal fracture. EMG helped us make the decision. The patient regained a normal vocal range.

Sonninen: Does EMG have any diagnostic value in cases with abnormal laryngeal sensations?

Kotby: EMG is not indicated because it is a sensory problem, and we would not expect to find any change in the motor system.

Loebell: Besides differential diagnosis, I want to stress that we must learn more about patients with central lesions and with so-called dysarthrophonia. EMG and neurography must be applied to as many of these patients as possible.

Kotby: Unfortunately, I am skeptical about what we will find in those cases. Other approaches, such as brain electrical activity mapping or evoked responses might be more productive.

Fauer: EMG may be better suited for clinicians in hospital practice than it is for clinicians in private practice. Logistics in private practice—where we usually have fewer assistants—make it difficult to use. This does not negate its usefulness in diagnosis or in prognosis but it suggests the clinical setting should be considered.

Fex: EMG is very useful for clinical diagnosis. But we are not very certain about its utility in prognosis, except in very special cases. It is important to perform several different EMG assessments of the same muscle. The motor units in the larynx are quite small, and therefore it is important to change the location of the electrodes so as to sample several of them.

Kotby: The problem is not as serious as it might seem. The concentric needle electrodes provide an average of the electrical activity in a spherical region, having a diameter of perhaps 3 mm. This volume will certainly include fibers from several different motor units. Thus, each electrode placement gives an overall picture of aberrant innervation across several motor units.

Hirano: Before we used multichannel recording, our procedure was to first examine either CT or CA on the normal side, then we inserted the needle into the paretic side, and compared the two. If there is any activity on the paretic side we record it, but to make sure of what we are getting we move the needle in or out a bit. If the EMG record is not changed by this maneuver, then we know we have a valid record. But if we initially get no muscle potentials, and moving the electrode fails to demonstrate any, then we know we have a total paralysis. In our latest series, 39 cases out of 130 showed complete paralysis.

Hirano: The reasons for the lack of popularity of laryngeal EMG are to be found primarily in the discomfort, both physical and psychological, associated with it. Medicolegal considerations also influence the decision to use EMG. Thus, in the United States, where medicolegal problems are quite significant, there is less use of clinical EMG than in Japan, where we have no legal threats.

Fex: Working with Bjorn Fritzell in Sweden, I questioned a large number of patients about their perception of discomfort associated with electromyographic examination. They reported very little discomfort, which is probably attributable to the very sharp and well-made needles that are now available. The cutaneous puncture is very easy, and once the skin is passed there is little sensation.

Hirano: Japanese patients report that the problems are largely psychological, and on this plane they feel EMG is a lot easier to deal with than fiberscopic examination.

Titze: I have been a subject several times now. I agree that the immediate discomfort of the fiberscope is greater, but I wonder about the lasting ef-

fects. It is true that the fiberscope irritates superficial tissues, but the invasion of tissue inherent in needle placement raises concern about possible scarring, especially when you change needle direction without removing it from the muscle.

Brewer: I still find the issue of interpretation of EMGs to be troublesome. We now know more about the location of the motor endplates in the muscles. But we cannot localize easily when we insert the electrode. We know that there is tremendous variation in the recordings, so that duplicating findings to the point of getting identical ones is all but impossible. Thus, when considering changes in interference patterns we are in a very slippery position. I am left feeling uncomfortable about it, especially in cases where we are trying to do a differential diagnosis of paresis versus myoparesis and hypofunction.

Kotby: I agree completely. The degree of muscle activity at a given moment may not differentiate neuropathology and myopathology. My feelings about the use of EMG at this moment are such that I am convinced that we must have accompanying and coordinated measures if we are to formulate a diagnosis.

Hirano: The fact is, though, that whenever we have a case in which we do an EMG examination we also have a lot of other information. So the comprehensive diagnosis may involve EMG, but its interpretation is tempered by the other data that are available.

Kotby: I could sum it up by saying that in cases of mobility disorders of the vocal fold, our augmented visual techniques cannot answer all the questions. So we need a tool to study the neuromuscular status of the vocal folds. Despite the fact that EMG is invasive, good preparation and explanations to the patient can eliminate a lot of the apprehension associated with EMG.

TERMINOLOGY

Laver: You can generate a thousand metaphorical labels, and they will be useful within a closed community. These would be of no utility from the scientific point of view. Totally descriptive labels, before they are supported by a community such as ours, must have a scientific basis. That is, there must be an objectively establishable link between the physiological cause, the acoustic consequence, and the auditory label that is assigned. Now, that is a counsel of perfection, because the necessary acoustic and physiologic research still lies perhaps 50 or 100 years ahead of us. If this group sets up a subcommittee to work on terminological questions, then that subcommittee should address itself to the principles on which labels are constructed, as well as to the labels themselves. My own personal preference as a theoretical phonetician is to have componential labels for those elements of the auditory complex that can be separated out linked to those elements of the physiological complex that can be separated out. The great power of a componential labeling system is that with a small number of labels you can generate a large number of output descriptors. So I would ask for any effort that we might undertake in terminology to address issues of principles, as well as labels themselves.

Sonninen: Although I have invested many years in developing useful terminology I am not capable of accomplishing this single-handedly. We seem to agree that the effort is needed and important. Therefore, I propose the creation of a multilingual terminology subcommittee to work on terminological issues via the mails.

Baken: It is too late to deal with terms that have been in circulation for many years. We all have our biases about them, and they are not going to change. But as we deal with an evolving technology, and an evolving view of vocal production, its assessment, and its diseases, we have an opportunity to influence the terms that are used as they are born. As we do that the group that makes initial recommendations should be a multilingual group. As we decide on terminologies, as we decide on what is good and bad, one of our concerns must be that a term be easily translatable into the languages that are spoken by a majority of the professionals in our fields. That needs to be considered ab initio.