882

INSTRUMENTS AND TECHNIQUES FOR INVESTIGATING NASALIZATION AND VELOPHARYNGEAL FUNCTION IN THE LABORATORY: AN INTRODUCTION

RENA A. KRAKOW*
***MARIE K. HUFFMAN*†**

* *Department of Speech-Language-Hearing*
Temple University
Philadelphia, Pennsylvania 19122
and
Haskins Laboratories,
New Haven, Connecticut 06511

† *Eloquent Technology, Inc.*
Ithaca, New York 14850

1. INTRODUCTION

Speech research in the laboratory can take many forms depending, in part, on the level or levels of speech organization one wishes to study. As illustrated by the contents of this volume, one can study the activity of muscles controlling articulator movement, the movements themselves, the aerodynamic conditions set up by the various movements, their acoustic consequences, and the impact of the acoustic signal on the listener. This article describes a variety of instruments and techniques that can be used to study nasalization at various levels in the speech production process. It is intended to aid in the interpretation of theoretical findings in this book (and elsewhere) in light of the instruments and techniques used, and to encourage laboratory research on nasalization and velopharyngeal function. Laboratory research provides insight into phonetic bases for phonological patterns

Phonetics and Phonology, Volume 5

involving nasalization, as well as an understanding of the nature of both normal and disordered patterns of velic motor control.

The focus here is on methods for studying nasalization with physiological, aerodynamic, and acoustic data; that is, speech production research. (For discussion of research on the perception of nasality, the reader is referred to the articles by Beddor and by Kurowski & Blumstein, in this volume.) In research of this kind, the choice of experimental methodology depends not only on the level(s) of speech production one wishes to study, but also on such factors as invasiveness, discomfort or risk to the subject, ease of use, expense, and inherent limitations in the instrumentation or available analysis methods. Such issues are addressed in our discussion of individual techniques; however, it is appropriate to mention here a few general considerations.

Some of the techniques we describe are invasive (e.g., endoscopy, intramuscular electromyography), and this may affect a subject's willingness to undergo a procedure, particularly in a nonclinical situation. Although some of these techniques can also be uncomfortable, in most cases the discomfort can be reduced (or eliminated) with the use of a topical anesthetic. Nonetheless, an invasive technique would likely be ruled out in studies of children, except when carried out for clinical purposes (e.g., as related to cleft palate). Another consideration affecting the use of certain invasive techniques may be the availability of a physician who can assist with the experiment. Such assistance can help to ensure proper insertion and removal of such devices as the Velotrace, fiberscope, or intramuscular electrodes, as well as appropriate application of some types of anesthetics used in these procedures.

Of the techniques described in this chapter, only X rays pose a known health risk, and this risk has been drastically reduced with the development of X-ray microbeam technology. But questions have been raised about other techniques (e.g., magnetic resonance imaging and electromagnetic articulometry), although the risks seem minimal, relative to the risks of ordinary day-to-day living. Nonetheless, subjects must be fully informed of any concerns before they consent to participate in such experiments.

It is important to bear in mind that the techniques alone do not guarantee a high-quality study. Experimental techniques are valuable only in the context of a well-designed study that includes careful analysis procedures. While it is beyond the scope of this chapter to examine such issues in detail, we do want to emphasize the importance of including proper control sequences when investigating the nature of nasal segments and their influence on the surrounding context. For example, velic movements in sequences containing a nasal consonant (e.g., *bean*) must be compared with movements in minimally contrastive sequences containing only oral segments (e.g., *bead*), in order to separate the contextual effects of the nasal consonant from intrinsic velic positions for the adjacent string. (See Bell-Berti, 1980; Bell-Berti & Krakow, 1991, for further discussion.) As described elsewhere in this book, measures of velopharnygeal activity (e.g., velum movement, lateral pharyngeal wall movement, velopharyngeal port aperture) show systematic differences among, as well as between, oral and nasal speech sounds. The techniques described here are useful for studying the variety of segmental as well as nonsegmental (e.g., stress, syllable structure, speaking rate) influences on velopharnygeal activity.

Although this chapter focuses on research methods that can be used to study nasalization and velopharnygeal function, we cannot overemphasize the importance of obtaining such data along with data on other articulatory subsystems (or their acoustic or aerodynamic consequences). Thus, for example, data on velic movements along with lip, jaw, and tongue movements provide invaluable information about speech motor control and coordination. As described below, some techniques are more conducive to the combined study of different articulatory subsystems than are others, and this is often an important consideration in an experimenter's choice of instrumentation.

In the sections that follow, our discussion of techniques and instrumentation is organized according to the level of speech production investigated. We discuss: (1) methods for studying sources of movement in the form of muscle activity, (2) methods for studying movement patterns, and (3) methods for studying the aerodynamic and acoustic consequences of movements. Within these three groupings, each technique or instrument is described individually. However, it is often possible, and clearly desirable, to acquire information on multiple levels of speech production simultaneously so that muscle activity can be related to movement patterns, and movement patterns to their acoustic and aerodynamic consequences.

In our discussion of movement sources and patterns, the emphasis is on how particular instruments can be used to monitor velum movement, lateral pharyngeal wall movement, velopharyngeal port aperture, and the muscle activity that controls these adjustments. In our discussion of movement effects (aerodynamic and acoustic consequences of velopharyngeal adjustments), on the other hand, the emphasis is on techniques for extracting information on velopharyngeal function from signals that reflect more global properties of the vocal tract.

In the last section of this chapter, we address general considerations governing the acquisition, storage, and manipulation of experimental data. In all sections, we provide references to other books, chapters, and articles that should be consulted by readers who wish to go beyond the scope of what is provided here.

2. MOVEMENT SOURCES: ELECTROMYOGRAPHY

Investigations of muscle activity have been important in studies of nasalization because they provide information on the control of velopharyngeal aperture. In normal speakers, the positions of the velum and lateral pharyngeal walls determine whether the velopharyngeal port is effectively open (allowing aerodynamic and acoustic coupling with the nasal cavities) or closed (preventing such coupling). Velopharyngeal port closure is accomplished by superior–posterior movement of the velum to make contact with the posterior pharyngeal wall and medial movement of the lateral pharyngeal walls to make contact with the lateral margins of the velum. (For detailed discussions of anatomy and physiology of the velopharyngeal region, see Dickson & Maue-Dickson, 1980; Folkins & Kuehn, 1982; Kahane, 1982). Data have been gathered on a number of muscles in the vicinity of the velopharyngeal port to determine which muscles control the movements of the velum and pharyngeal walls (see Bell-Berti, this volume). These studies involve a technique known as *electromyography* (EMG). (For a detailed discussion of EMG and muscle activity, see Basmajian, 1974.)

Electromyography makes it possible to detect and record the time-varying electrical activity associated with muscle contraction. Although the earliest use of this technique was for medical purposes, since the 1950s numerous studies have used EMG techniques for examining a range of muscles active in speech, including those that control velopharyngeal port aperture (e.g., Bell-Berti, 1976; Draper, Ladefoged, & Whitteridge, 1959; Fritzell, 1969; Henderson, 1984; Hirano & Ohala, 1969). Understanding the technique requires some understanding of muscle function itself.

The minimal unit of muscle activation, the *motor unit,* consists of a motor neuron and the muscle fibers that it innervates. When an impulse is transmitted to the interface between nerve and muscle (the *motor endplate*), a wave of electrical activity passes along each of the fibers in the motor unit. Electrodes placed in the vicinity of the disturbance can detect and record this time-varying electrical activity. The raw signal obtained in most EMG studies of speech is an interference pattern, or the summed pattern of activity of many motor units. Generally, the motor units that are sampled represent a subset of those that comprise the muscle.

Typically, EMG electrodes are used in pairs; the experimenter obtains information on the electrical activity at two points (the sites of the electrodes) and compares the activity at those points over time. Wider spacing between the two electrodes results in a larger recording field. A differential amplifier is used to amplify the difference between the electrical activity transduced by the two electrodes. Because the electrical signal of the muscle is inherently weak (generally on the order of a few hundred microvolts), it must be amplified; furthermore, because the signal is subject to interference from ambient electrical activity, it is important to use the shortest possible leads between the electrodes and the amplifier. For analysis purposes, the signal is often rectified (converted to all positive values) and smoothed by means of integration. In this form, the signal provides an indication of the relative strength of muscle activity over time (see Harris, 1981).

EMG electrodes can be of two basic types: surface (suction, glue-on, or paint-on) or intramuscular (conventional needle or hooked-wire). Procedures that use surface electrodes have the advantage of being noninvasive, but their use is restricted to muscles that are near the skin or mucosal surface and are in a well-defined region, e.g., muscles of lip protrusion. Electrical activity farther from the surface will likely be too weak to be detected at the surface, and when in the vicinity of overlapping muscles, surface electrodes may pick up potentials from any or all of them, leaving the researcher uncertain as to the specific muscle(s) that are acting. (For additional information on surface electrodes, see Abbs & Watkin, 1976; Cooper, 1965; Harris, 1970). Surface electrodes have not been able to supply data on specific muscles that control velopharyngeal function because of the overlapping arrangement and depth of some of the muscles in this region. (See Harris, Lysaught, & Shvey, 1965, and Lubker, 1968, for studies showing the kinds of data obtained with surface electrodes placed in the velopharyngeal region.)

In contrast to surface electrodes, intramuscular electrodes are inserted directly into the muscle(s) of interest and can therefore be used to study the activity of any muscle. In early studies, needle electrodes were used, but their rigidity and size made their use in the oral cavity during speech extremely problematic. The rapid movements that occur in speech often caused the needle electrodes to move and even dislodge during an experiment, resulting in movement artifacts in the recorded signal and considerable discomfort for the subject.

More recent studies have used hooked-wire electrodes that can be inserted perorally into most palatal and pharyngeal muscles. They are thinner and lighter than needle electrodes, which makes them more flexible, and thus more likely to remain in position during speech and more comfortable for the subject. Hooked-wire electrodes can be made by passing a doubled strand of fine insulated wire through a hypodermic needle to form a loop (Figure 1). The insulation is then removed from the end of the loop and the loop is cut to form two bare ends of wire, one extending slightly farther than the other. The bare ends are then bent back to form hooks. The two hooks must be of different lengths to avoid contact between their ends in the muscle, which would result in a short circuit. After the application of a topical anesthetic, the needle is inserted into the muscle and withdrawn, leaving the ends of the wires embedded (or hooked) in the muscle (Basmajian & Stecko, 1962). The electrodes are then attached to the differential amplifier via the external ends of the wires. (At the end of an EMG study using hooked-wire electrodes, a slight tug on the wires removes them from the muscle).

After inserting the electrodes, it is necessary to verify their placement. For the levator veli palatini, which contracts to raise the velum and is considered the most important muscle in velopharyngeal activity, the electrode-bearing needle is inserted (perorally) into the velar "dimple" as the subject sustains the vowel [ɑ]. The needle is directed roughly 10 mm from the surface in a latero-cranio-posterior

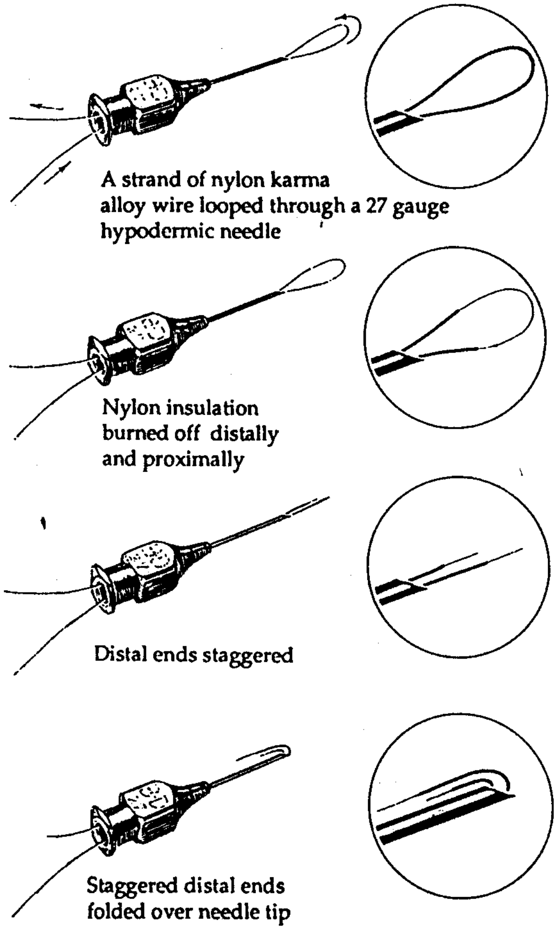

Figure 1. Schematic showing the construction of hooked-wire electrodes. Adapted from Basmajian and Stecko (1962) with permission.

direction (Hirose, 1971). To verify electrode placement in the levator palatini (which contracts to raise the velum), the subject is asked to produce a sustained [s]. The levator is expected to show marked activity associated with the strong raising gesture of the velum necessary to ensure adequate oral air pressure for this consonant. Obviously, verification gestures are problematic when the function of the muscle under investigation is unknown or not well understood. Researchers must then rely on known anatomical landmarks, yet they must also remain aware that there are individual differences in the precise arrangement of muscles (see Harris, 1970, and Hirose, 1971).

Numerous studies have employed hooked-wire electrodes to determine which muscles influence velopharyngeal function and in what manner. As noted above, such studies find a systematic relation between levator palatini activity and velum movements. Contraction of the levator palatini is observed prior to velum raising, while relaxation of the muscle is observed prior to velum lowering in strings of oral and nasal segments (Figure 2). In addition to levator palatini, the muscles investigated for their role in velopharyngeal function using EMG include palatoglossus, palatopharyngeus, superior pharyngeal constrictor, and middle pharyn-

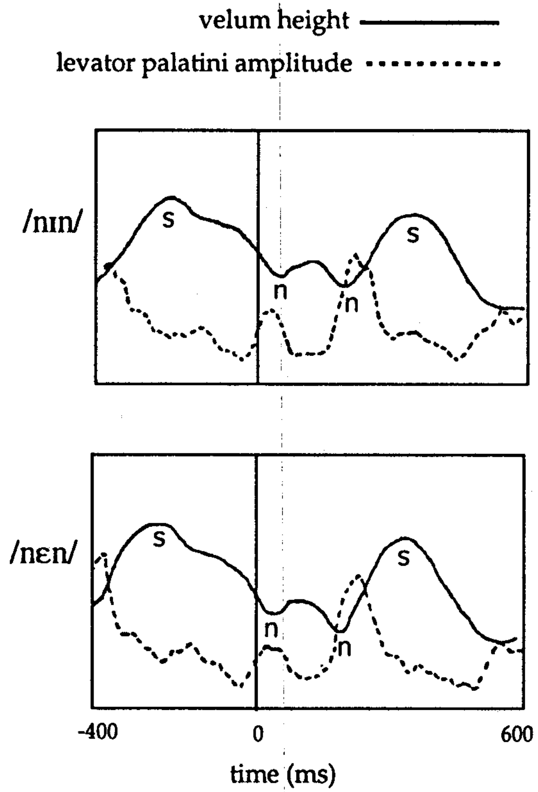

Figure 2. EMG data obtained with hooked-wire electrodes from the levator palatini (dashed lines) along with the corresponding vertical velum movement data obtained with a fiberscope (solid lines). The target items are /nɪn/ and /nɛn/, and the carrier phrase is *Guess* ____ *soon.* The movement data show velic peaks for the /s/'s of the carrier phrase and valleys for the /n/'s. Between the /n/'s, a small peak is seen for the contextually nasalized vowel, which has a higher velic position than the nasal consonants but a lower position than the oral consonants. (The symbols *s* and *n* are used to show the velic target positions for these sounds.) Correspondingly, the EMG signal shows peaks preceding velic raising for the /s/'s and valleys (corresponding to relaxation) preceding velic lowering for the /n/'s, as well as a small peak preceding the vowels in between. Adapted from Henderson (1984) with permission.

geal constrictor. Information about the activity of these muscles can be found in Bell-Berti (1976, this volume), Dickson and Maue-Dickson (1980), Fritzell (1969, 1979), and Henderson (1984). For a more complete discussion of EMG, placement of electrodes for studying velopharyngeal function, and verification gestures, see Harris (1981) and Hirose (1971).

EMG data can provide important information on articulatory coordination and control because it is possible to study a number of muscles at one time, whether associated with the same or different articulators. Furthermore, EMG data can be collected simultaneously with other kinds of data, such as movement data or acoustic data. Sampling the set of muscles affecting a single region along with the corresponding movement data is particularly useful for determining the relation between muscle activity and movement (see Henderson, 1984), which is not always as straightforward as one might expect. For example, at times a given muscle may contract to augment the effect of another muscle on the same articulator or to counteract its effect.

3. MOVEMENTS

Movement data on the velum and pharyngeal walls can be gathered with a number of different instruments. In general, techniques used to study movement patterns of the articulators can be divided into two basic types. One type produces an image or picture of a region of the vocal tract. Imaging techniques provide information on the shape as well as the position of articulatory structures; some imaging techniques provide only static views, while others can be used to obtain dynamic information by repeated or continuous imaging. The other type of technique for studying articulatory movements is inherently dynamic but provides position, rather than shape, information. This type involves tracking the time-varying position of a point on an articulator, i.e., a *fleshpoint* (e.g., on the tongue, lips, jaw, or velum) (in the vertical and/or horizontal dimension) or the time-varying magnitude of opening in a region that opens and closes (e.g., velopharyngeal port, glottis, lips).

3.1. Imaging Techniques

3.1.1. RADIOGRAPHY

Radiography has provided a large body of information concerning articulatory activity in speech, including velopharyngeal function. When an X ray is taken, photoelectron beams are passed through the body, with the source on one side of the subject and a receiver on the other. Because tissue density affects the extent to which the beams are absorbed by the body, the output reflects the varying tissue types in the X-ray path. The electron beam that passes through the subject can be converted to an image of relative tissue density. In speech research, the images are usually taken in lateral view.

Single-frame X rays provide detailed images of the vocal tract, which make it possible to measure static articulator positions relative to fixed bony structures (Buck, 1954; Hixon, 1949; Subtelny, 1956). The X-ray energy emerging from the subject is imprinted directly on film or is intensified (on the order of 1000–3000%) and projected to a fluoroscope. The fluoroscope screen is photographed for a permanent record of the image. Illumination of the fluoroscope requires less incident X radiation than does film, so the X-ray energy absorbed by the subject and consequently the inherent risk is reduced by the fluoroscopic technique. (However, images imprinted directly on film tend to have higher resolution than those filmed from a fluoroscope.)

Dynamic information can also be obtained with X-ray images that have been intensified and projected on a fluoroscope, by filming (cinefluorography) the projected images at high speed (up to 150 frames/sec or higher). This technique is known as cinefluorography. As soon as the images have been captured on film, they can be projected onto paper, frame by frame, and the articulatory structures of interest along with bony reference structures can be traced. Cross-frame analyses to obtain the time-varying articulator movements are made by comparing articulator positions when the bony structures are aligned (Moll, 1960). Measures of velum height and of velopharyngeal aperture were obtained in a cinefluorographic study by Kent and Moll (1969), along with measures of other articulators, including the lips, jaw, tongue, pharynx, and hyoid bone. Figure 3 is a tracing from this study showing how articulator positions were measured with reference to the hard palate.

The development of motion picture X rays was a great advance from single-frame techniques because it enabled researchers to study speech dynamically. Nonetheless, there remain a number of problems. One of the drawbacks of imaging with X rays (whether still or motion picture) is that the X-ray beam passes through everything in its path, often yielding images of structures on top of other structures. Because denser tissues, e.g., bones, produce darker images, they can obscure less dense softer tissues (e.g., the tongue or velum) in the same path. To increase soft tissue visualization during cine- or videofluorography, some researchers place a contrast medium such as barium sulfate paste (e.g., McClean, 1973) or a thin gold chain (e.g., Alfonso & Baer, 1982) on the articulatory structure(s) of interest (e.g., Kent & Moll, 1969; McClean, 1973).

A more serious concern about using cinefluorographic techniques is the risk to subjects. Naturally, since X rays taken over time result in an increase in the cumulative radiation dose relative to single-frame X rays, the risk to the subject also increases. Because of the risk, such experiments must be quite limited in duration. Hence, the number of tokens and/or utterances that can be collected is also

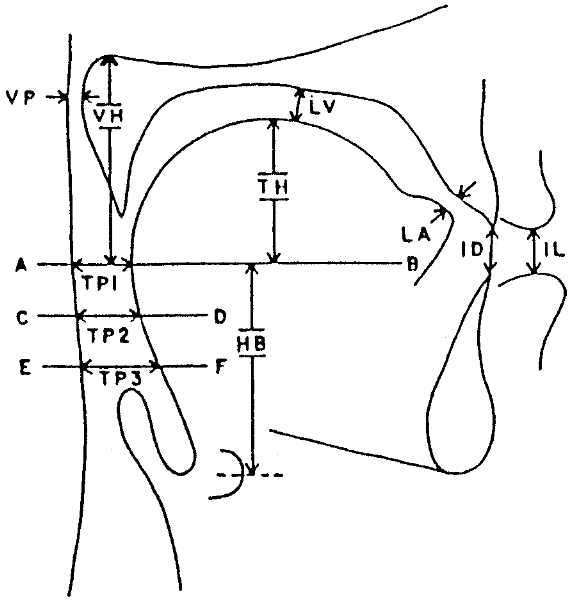

Figure 3. Tracing of a lateral view X-ray cinefluorographic frame showing measurements of a variety of articulatory positions: ID (interdental distance), IL (interlabial distance), LA (lingua–alveolar distance), LV (lingua–velar or lingua–palatal distance), TH (tongue height), VH (velar height), VP (velopharyngeal aperture), TP1, TP2, TP3 (measures of pharyngeal width), HB (hyoid bone position). AB, CD, and EF are reference lines drawn parallel to a line tangent to the hard palate. Adapted from Kent and Moll (1969) with permission.

limited. Further limiting the size and scope of these studies is the fact that frame-by-frame analysis is extremely time-consuming.[1] Nonetheless, a number of studies using this methodology have provided important information about velopharyngeal function (see, e.g., Kent & Moll, 1969; McClean, 1973; Moll & Daniloff, 1971).

It should be noted that in recent studies some researchers have obtained videotapes (videofluorography) rather than cine films (cinefluorography) of X-ray images over time. Probably the main reasons for choosing video over cine are that less radiation is required to record the image on video (Perkell, 1977) and video images are immediately available (i.e., they do not have to be developed). On the other hand, standard videotape recording systems have frame rate limits of 50–60/sec, whereas motion picture cine films can be made at rates as high as 150 frames/sec or higher. Also, the resolution achieved on film is generally better than on video. Generally speaking, however, the problems of radiation risk and of laborious analysis are true for both video- and cinefluorographic studies. (Radiographic techniques that involve fleshpoint tracking are discussed in Section 3.2.)

3.1.2. MAGNETIC RESONANCE IMAGING

Magnetic resonance imaging (MRI) is a relatively new technique that can provide detailed three-dimensional information on articulatory structures and airways. MRI produces tomographic images (i.e., a series of images of thin sections) of human tissue in sagittal, coronal, oblique, or transverse planes. Although X-ray techniques have been adapted to provide tomographic images in the form of computed tomography (CT) scans, the radiation risk involved precludes their use for basic research in speech.[2] In contrast to X-ray techniques, MRI involves no ionizing radiation and also provides excellent soft tissue discrimination. Although developed for medical imaging purposes, MRI has recently been used in a number of studies of vocal tract shape for different speech sounds (Baer, Gore, Boyce, & Nye, 1987; Baer, Gore, Gracco, & Nye, 1991). To date, the use of MRI in studies of velopharyngeal function has been very limited, but because the technique holds promise for future studies, we describe it here.

MRI images represent the relative densities of hydrogen nuclei in tissue. Hydrogen nuclei are abundant in water, and because different tissues differ in their water content, the images show differences among tissue types. A subject is placed in a strong magnetic field that causes the hydrogen nuclei in the body to align along the direction of the magnetic field's poles. Radio frequency impulses are then applied at the resonant frequency of the hydrogen nuclei, temporarily knocking them out of alignment with the magnetic field. As they realign, the nuclei emit their own radio signals. What MRI captures is the time to realignment. Realignment time, and hence the nature of the signal emitted, is a function of hydrogen density and therefore of tissue type.

MRI involves no ionizing radiation (and is therefore much safer than X-ray techniques), it is noninvasive, and it lends itself to efficient analysis procedures. Each image is stored directly on the computer as a matrix of numbers that provide a high-resolution display in the form of pixels in grayscale. Given the possibility of multidimensional imaging along with the advantages of computer-based processing and the relative safety of MRI, this technique is considered the best available for obtaining data on static vocal tract shapes (see Baer et al., 1991). However, the use of MRI in speech research is limited by a number of factors: First, with the MRI units currently available, subjects must be in a supine position, a common position for medical imaging but an uncommon position for speech production. The effects of gravity on speech structures in this position must therefore be evaluated (see Baer et al., 1987; Whalen, 1990). (Actually, MRI units could be constructed so that a subject could sit or stand; current units, however, were designed for imaging with patients who are unable to do so.) Second, the amount of time necessary to obtain a clear image with MRI can range from several tens of seconds to several minutes, limiting investigation to static vocal tract shapes maintained for unnatural lengths of time. For example, in the studies of Baer et al.

(1987, 1991), subjects produced sustained vowels, taking care to hold the oral tract configuration between inspirations occurring at about 15-sec intervals. However, the imaging times have been decreasing with newly developed MRI units, so we can probably expect to see MRI used in the future for more dynamic sorts of measures.

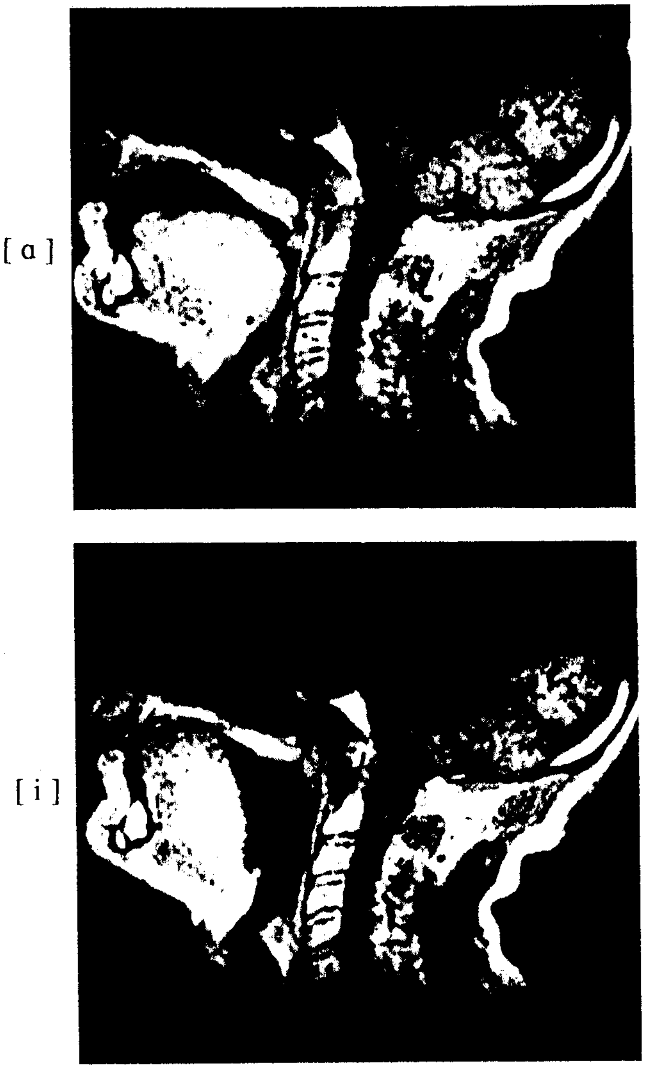

Figure 4. MRI images obtained from an adult male speaker producing the steady-state vowels [ɑ] and [i]. Figure courtesy of Patrick Nye.

To date, study of the velum or velic region with MRI has been extremely limited. This reflects not only the fact that MRI is a relatively new technique in the study of speech but also specific concerns about the effects of gravity on the velum (see Whalen, 1990). That the technique might be useful for such study is suggested by the differences in velum position for low and high vowels that can be seen in MRI images such as the mid-sagittal sections in Figure 4 (see also Whalen, 1990). Although the velum is backed against the posterior pharyngeal wall for both the high and low isolated vowels in these images, it is clear that the configuration of the velum differs in the two cases. The velum is higher for [i], with an apparently tighter constriction of the port than for [ɑ]. Because many studies of American English show some port opening for low vowels such as [ɑ], we must consider the possibility that the apparently closed port shown for this vowel in the MRI image may be a function of the subject's supine position. Nonetheless, the image shows that there are differences between the two vowels and enables us to see the shapes of other structures simultaneously. As noted earlier, MRI units are being developed that can create images more rapidly, so that it ought to be possible to obtain dynamic measures in the future and, as noted, units could be constructed to allow the subject to sit or stand.

3.1.3. ULTRASOUND

As with radiography and MRI, ultrasound makes it possible to obtain images of structures inside (as well as outside) the vocal tract without passing an instrument inside. In speech research, ultrasound has been particularly useful in providing time-varying position and shape information on the tongue (Stone, Shawker, Talbot, & Rich, 1988; Shawker, Sonies, & Stone, 1984). It has also added to our understanding of how velopharyngeal port aperture is controlled by providing information on lateral pharyngeal wall movement (Kelsey, Minific, & Hixon, 1969; Minifie, Hixon, Kelsey, & Woodhouse, 1970; Parush, 1984; Ryan & Hawkins, 1976).

Ultrasound is so named because it employs high-frequency sound waves out of the range of human hearing. (In clinical studies or the sorts of research studies we will address, the frequency used is usually above 1 MHz). With an ultrasonic transducer placed against the skin, the sound waves penetrate the various structural layers as they move through the body.[3] Because these layers differ in acoustic impedance, some of the sound is reflected back toward the source at the boundaries between tissue types. The tissue–air boundary is the most reflective boundary, causing an almost complete reflection of the ultrasonic waves. Ultrasound is, therefore, particularly useful for studying those structures adjacent to air, such as the upper surface of the tongue, the vocal folds, and the lateral pharyngeal walls.

The ultrasonic sound waves are generated by electrically exciting a piezoelectric crystal that deforms and transduces the electric energy into acoustic energy. In *pulse-echo ultrasound* (used in many of the speech studies), the crystal used to generate the sound waves also receives the *echo* (the reflected sound) and transduces the acoustic energy back into electrical energy, which can then be stored and measured. The signal is transmitted by exciting the crystal for brief periods at a particular rate. Measurements of the time of arrival of the reflected sound waves provide information about the distance between the externally placed transducer and the various tissue boundaries in the path of the ultrasound signal. If the anatomy of the region is well known, the echoes can be associated with specific structures in the ultrasonic path. Such information obtained over time provides dynamic information on articulator position and/or shape.

In studies of lateral pharyngeal wall movement, the technique essentially tracks a point (i.e., that point in the path of the ultrasound signal) on the lateral pharyngeal wall over time. In that sense, it would be more appropriate to discuss ultrasound in Section 3.2 on fleshpoint tracking. However, ultrasound is used more commonly as an imaging technique (e.g., in diagnostic imaging and in studies of the tongue by speech researchers), and the images are built from repeated scans of a region, composed of the reflections from individual fleshpoints in a sector of tissue. It is useful to understand both kinds of ultrasound techniques and easier to explain the imaging technique by first explaining the fleshpoint-tracking technique; thus, both are included in this section. We begin with fleshpoint tracking as used in studies of lateral pharyngeal wall movements.

Figure 5a shows the placement of an ultrasound transducer against the external neck wall beneath the earlobe and behind the ramus of the mandible. The transducer is directed medially towards the ipsilateral pharyngeal wall at the level of the velopharyngeal port (see Parush, 1984; Ryan & Hawkins, 1976). Determination of correct positioning for detecting lateral pharyngeal wall movement is accomplished by having the subject swallow, which causes a strong medial gesture of the pharyngeal wall that should be visible in the ultrasonic signal. With the transducer positioned to detect lateral pharyngeal wall movement, the echo of the moving tissue–air interface can be viewed in the form of a time–amplitude display, where the amplitude of the signal indicates relative distance of the tissue–air interface from the transducer (Figure 5b). Although one may pick up ultrasound echoes due to tissue discontinuities that occur between the external neck wall and the lateral pharyngeal wall, it is possible to adjust the gain so that only the tissue–air interface (which should provide the strongest signal) is detected. Detailed discussion of the use of ultrasound in studying lateral pharyngeal wall movement can be found in Parush (1984).

To obtain an image of a sector (a thin section or slice) of tissue, such as a section of the tongue, the transducer is rotated against the skin (or several transducers rotate around a single axis), and the gain is set so as to obtain reflections from tissue discontinuities that lie between the transducer and the tissue–air interface. When the tongue is imaged in this way, the surface (tissue–air boundary) and various soft tissue structures within the tongue, such as the genioglossus, the floor of the intermuscular septum (which separates the tongue and jaw muscles), the geniohyoid, and the mylohyoid can be distinguished. Repeated imaging of this kind can produce a motion picture display of the region showing the changing configuration of the structures over time (Stone et al., 1988).

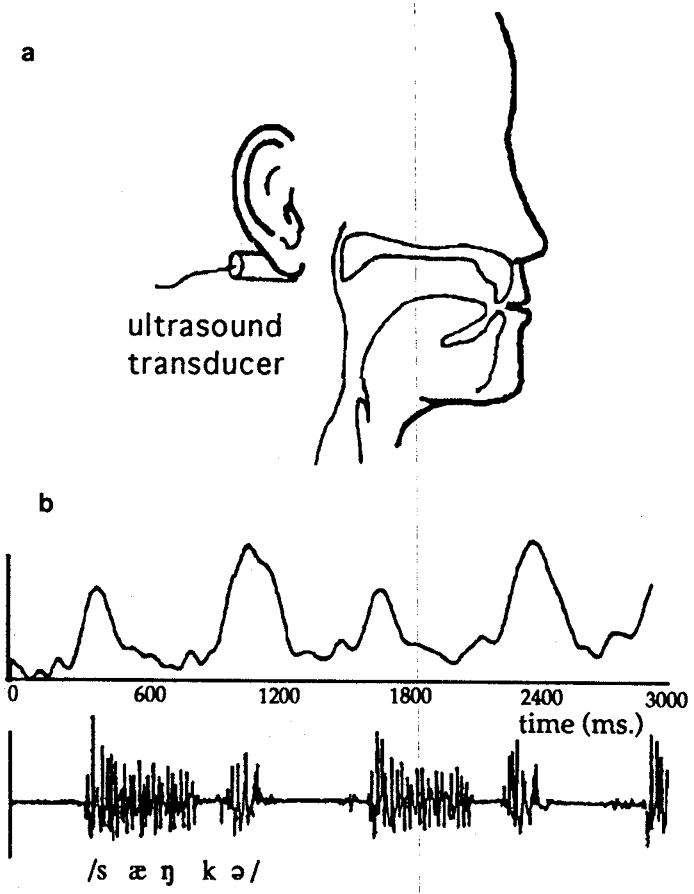

Figure 5. (a), Schematic showing the positioning of an ultrasound transducer for obtaining measures of lateral pharyngeal wall movement. (b), Ultrasound signal showing lateral pharyngeal wall movement and corresponding acoustic signal for one speaker's productions of two repetitions of the word *sanka*. The upper peaks in the ultrasound trace indicate more medial positions of the pharyngeal wall (a greater distance between the transducer and the ipsilateral pharyngeal wall), consistent with velopharyngeal port closure. Adapted from Parush (1984) with permission.

Ultrasound involves no discomfort for the subject and minimal, if any, hazard of exposure. In addition, data collection is rapid, the equipment is less expensive than X-ray or MRI imaging equipment, and some ultrasound units are portable.

The nature of the technique clearly limits what one can view, however, since the signal will not pass through a tissue–air boundary. Also, most of the ultrasound signal is reflected when it encounters hard tissue; thus, hard tissue structures (e.g., the hard palate) cannot be used for reference information.

3.1.4. FIBEROPTICS

Flexible fiberoptic endoscopes have been used in a large number of studies to obtain images of structures inside the vocal tract. In contrast to the techniques of MRI, X ray, and ultrasound, endoscopy involves inserting a portion (a flexible bundle of glass fibers) of a device (the fiberscope) into the subject's nose (or mouth in some cases, for laryngeal views). In speech research, fiberoptic endoscopes (fiberscopes) were first used to view the larynx, but they have since been used in numerous studies to examine the velopharyngeal mechanism (Sawashima & Ushijima, 1971, Sawashima, 1977).

The flexible fiberscope contains two bundles of thin, flexible glass fibers. One bundle serves as a light guide, carrying light from an externally attached light source into the region of interest. The other is a coherent bundle of fibers attached at the internal end to an objective lens, and it carries the image from the lens within the subject's vocal tract into an external eyepiece. The investigator can either view the region of interest through the eyepiece or attach a special adapter to the eyepiece, from which cine films or videotapes of the view can be made.

The thin flexible fibers of the two fiberoptic bundles are attached to a directional control unit that helps to pass the bundles through the winding passageways of the vocal tract. A fiberscope with a wide-angle lens can be used to view the lateral and posterior pharyngeal walls, the nasal surface of the velum, and the velopharyngeal port, while a narrow lens can be used to view the upper surface of the velum or the larynx. The different views require positional adjustments of the device (Sawashima, 1971, 1977) (Figure 6).

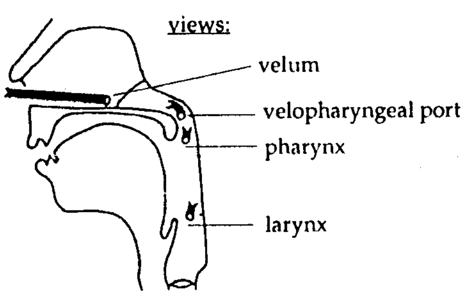

Figure 6. Schematic showing positions of the internal end of a fiberoptic endoscope for obtaining different views of the internal structures of the vocal tract. Reproduced from Sawashima (1977), with permission.

Figure 7a–b provides images (in the form of photographs and tracings) obtained with a fiberscope inserted through the nose and positioned to view the nasal surface of the velum. A clear difference in velum height between the oral vowel [e] (Figure 7a) and the obstruent [s] (Figure 7b) can be seen, with more extreme raising for [s] than for [e]. Figure 8a–b provides images obtained with the fiberscope positioned to view the velopharyngeal port, showing differences in port aperture for the low and mid vowels, [ɑ] (Figure 8a) and [e] (Figure 8b). Opening of the port is seen for both "oral" vowels, but considerably more for [ɑ] than for [e].

Although invasive, fiberoptic endoscopy is designed to allow researchers to view internal structures without radiation of any kind or any other health risk, assuming proper insertion of the device. Typically, any discomfort can be reduced or eliminated simply with the use of a topical anesthetic. Lacking the risk inherent in X rays, this technique allows for longer experiments (and therefore more data). However, as with X-ray images, frame-by-frame analysis is necessary with fiberscope images, which makes the analysis tedious.

Calibrating data obtained with the fiberscope can be complicated because the size of the image changes as a function of the distance between the structure viewed and the tip of the optical fibers. In some studies of velum height, researchers (e.g., Bell-Berti, Baer, Harris, & Niimi, 1979; Niimi, Bell-Berti, & Harris, 1982) insert a long, thin plastic strip with grid markings into the subject's nostril. The strip is positioned along the floor of the nose and the nasal surface of the velum. If the dimensions of the markings on the strip are known, then it is possible to correct for distortion in the view. Also, the markings on the strip serve to enhance the contrast between the edge of the superior velic surface and the posterior pharyngeal wall, making it easier to measure velum height.

Numerous studies have used a flexible fiberoptic endoscope to investigate velopharyngeal function, including Bell-Berti (1980), Bell-Berti et al. (1979), Bell-Berti and Hirose (1975), Benguerel, Hirose, Sawashima, and Ushijima (1977), Henderson (1984), and Ushijima and Sawashima (1972).

3.2. Fleshpoint-tracking and Aperture-tracking Techniques

Fleshpoint-tracking involves identification of points on an articulatory structure or structures of interest (e.g., lips, tongue, jaw, velum) and then monitoring the time-varying position of the point(s) in the horizontal and/or vertical dimensions. In aperture tracking, a value of relative opening of some region, such as the velopharyngeal port, glottis, or lips, is obtained over time. As compared with imaging techniques (described above), fleshpoint-tracking and aperture-tracking techniques produce data in a form which is much easier to manipulate.

We begin our discussion with radiography, which, although widely known as an imaging technique, can be adapted for point tracking. Most of the other techniques we discuss in this section are also point-tracking techniques. The last one

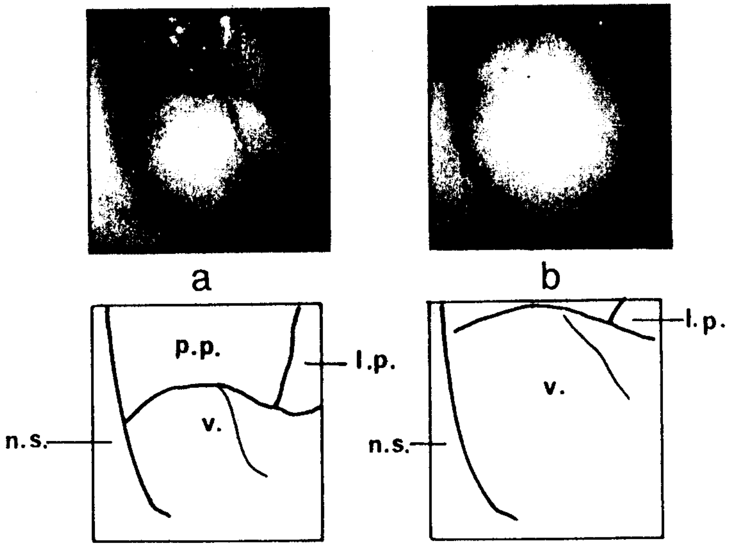

Figure 7. Photographs and their tracings obtained with a fiberscope positioned to view the velum in order to determine its relative height. Shown in (a) is [e] and in (b) is [s]. v. marks the velum; n.s. marks the nasal septum; l.p. marks the lateral pharyngeal wall; p.p. marks the posterior pharyngeal wall. Reproduced from Sawashima (1977) with permission.

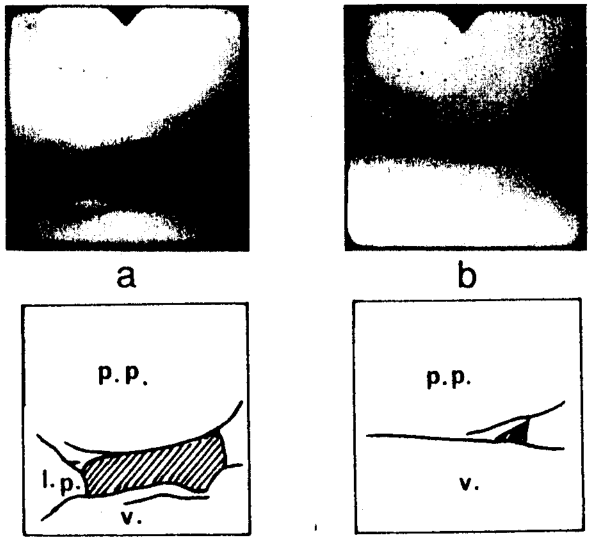

Figure 8. Photographs and their tracings obtained with a fiberscope positioned to view the velopharyngeal port. Shown in (a) is [a] and in (b) is [e]. v., l.p., and p.p. mark the velum, lateral pharyngeal wall, and posterior pharyngeal wall, respectively. Reproduced from Sawashima (1977) with permission.

we describe, photodetection, can be used for either point tracking or aperture tracking, depending on the design and positioning of the photodetection device.

3.2.1. Radiography

Fleshpoint tracking with X rays is accomplished with radiopaque markers (pellets) attached to the articulators of interest (e.g., tongue, lips, jaw, velum) in the midsagittal plane. The X-ray images of the vocal tract are intensified and projected onto a fluoroscope from which high-speed motion pictures are obtained. The time-varying positions of the radiopaque markers are then recorded over a series of X-ray frames, with fixed bony structures serving as references (see, e.g., Kent, Carney, & Severeid, 1974; Kuehn, 1976). The pellets are usually attached with a nontoxic adhesive to the articulators. However, because it can be difficult to get a pellet to adhere to the velum for any extended period of time, some researchers have sutured the pellet to the velum (e.g., Kent et al., 1974; Kuehn, 1976).

Figure 9a illustrates the location of pellets on the tongue and velum from a study by Kent et al. which examined velum, tongue, and jaw movement with fleshpoint tracking (1974). Figure 9b shows, for one of the subjects, the movement paths of the pellets on the velum and tongue during the course of the first syllable of the noun *contract.* An interesting observation made from these data about velum–tongue coordination was that the velum reached its minimum when the tongue was in a low position for the vowel /ɑ/ and that it began to rise well before the tongue achieved its target position for the /n/.

Building on the idea of pellet tracking, researchers at the University of Tokyo developed an X-ray system for studying articulator movements which both reduced the X-ray risk and enhanced the efficiency of analysis. The system was subsequently refined by researchers at the University of Wisconsin, and the facility in Japan closed. With the *X-ray microbeam system,* an X-ray beam is passed through a tiny pinhole in the direction of the subject (Figure 10). The narrowed beam is steered by a computer so as to scan those regions on the lips, tongue, jaw, and velum where pellets have been attached (with dental adhesive) in the midsagittal plane. The output of an X-ray detector is fed to a computer, which uses previous pellet positions to predict the region in which to direct subsequent scanning during speech. Initially, the entire area of the subject's vocal tract is scanned to determine the location of the pellets when the subject is at rest (i.e., in the initial position). The computer automatically stores the x–y coordinates of each pellet at the initial and subsequent positions. "Round-robin scanning" provides coordinate data on the positions of each of the pellets over time (see Abbs & Nadler, 1987, and Kiritani, Itoh, & Fujimura, 1975, for further discussion). The Wisconsin system was designed to be able to track pellets on the upper lip, lower lip, jaw, tongue tip, tongue blade, tongue body, tongue root, and velum at a rate of 125 samples/sec, along with pellets on the bridge of the nose and an upper tooth as reference points for detecting head movement (Abbs & Nadler, 1987).

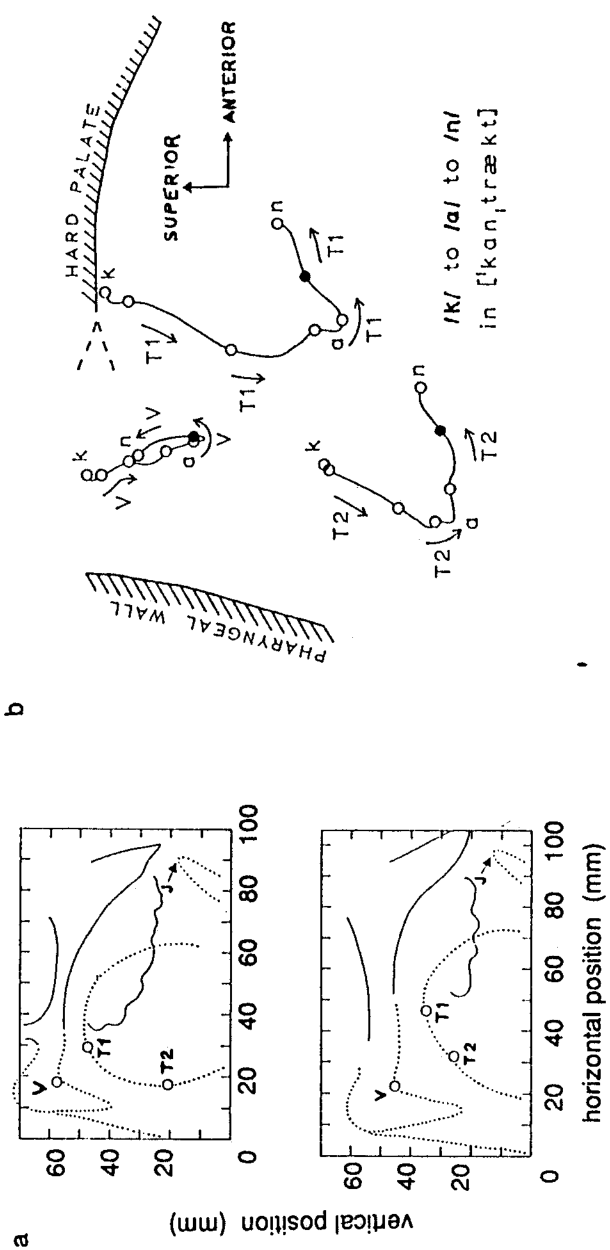

Figure 9. (a) The positions of velum (V) and tongue (T1 and T2) pellets are shown on lateral view X-ray tracings for two subjects superimposed on individualized coordinate grids showing bony reference structures. (b) Schematic showing the motion paths of the three (V, T1, and T2) fleshpoints for the [kan] portion of the noun *contract*. The arrows indicate the direction of fleshpoint movement, and the small circles mark fleshpoint positions sampled every 40 msec. The filled circle marks the time at which the velum begins its ascent. Adapted from Kent, Carney, and Severeid (1974) with permission.

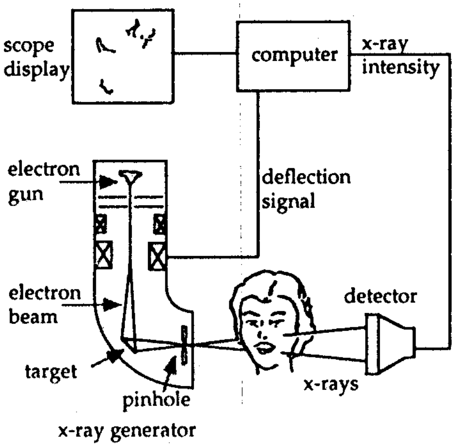

Figure 10. Schematic diagram of a computer-controlled X-ray microbeam system. Adapted from Fujimura, Miller, & Kirtani (1977) with permission.

Research using X-ray microbeam techniques to study velum movement patterns has been reported by Fujimura (1990), Fujimura, Miller, and Kiritani (1975, 1977), and Vaissière (1988). Because of problems affixing the velum pellet, these researchers glued the pellet to the end of a long plastic strip, which was inserted into the nose until the pellet was positioned on the surface of the velum.

An advantage of X-ray techniques, in general, is the possibility of obtaining data simultaneously on multiple articulators both inside and outside the vocal tract, allowing researchers to examine inter-articulator coordination. With the reduction in risk to subjects and an increase in the efficiency of analysis, the microbeam system made it possible to conduct larger scale studies than were previously possible.[4,5] Despite all the advantages of the X-ray microbeam system, however, the facility in Wisconsin like that in Japan has now been closed, largely because of the immense cost associated with this kind of equipment. Nonetheless, we can still expect to see a number of studies appearing in journals based on X-ray microbeam data because the data obtained in Japan have been archived and continue to be available to researchers; also, much of the recently collected data from Wisconsin have yet to be fully analyzed.

3.2.2. ELECTROMAGNETIC ARTICULOMETRY

Another technique for obtaining dynamic multi-articulator movement in the form of fleshpoint tracking is electromagnetic articulometry (or magnetometry). Magnetometry was first introduced as a method of articulator tracking by Hixon (1970), who devised a simple system for collecting data on jaw movement.[6] Recently developed systems, however, make it possible to obtain data on the time-varying positions of the lips, tongue, jaw, and velum in the midsagittal plane (Perkell, Cohen, Svirsky, Matthies, Garabieta, & Jackson, 1992; Schönle, Gräbe, Wenig, Höhne, Schrader, & Conrad, 1987).

In these newer systems, small transducer coils are attached to selected articulator fleshpoints in the midsagittal plane. The subject wears a head-mounted assembly which supports three transmitter coils. The transmitter coils produce alternating magnetic fields, which are intercepted by small receiver coils (transducers) glued to the articulators (Figure 11). The outputs of the receiver coils are fed to receiver electronics via fine twisted wires running from those coils. Because the magnetic fields emitted by the transmitter coils are weaker at greater distances from the transmitters, the transducer coils on the articulators intercept magnetic fields that vary in strength as a function of transmitter–transducer (i.e., transmitter–articulator) distance. The output voltages generated by the receiver electronics are digitized and converted to x- and y-coordinate values as a function of time, with specialized software.

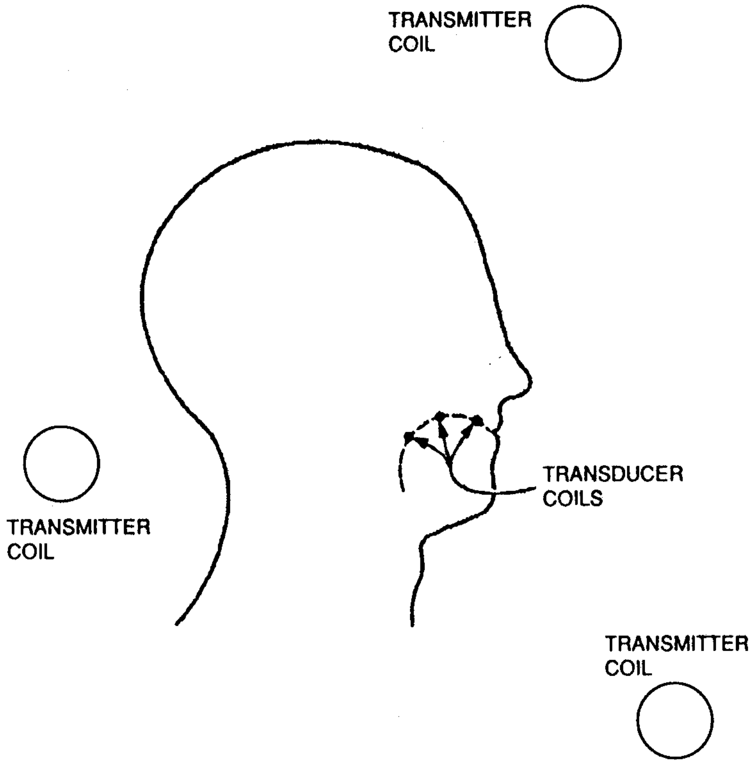

Figure 11. Schematic diagram showing the locations of transmitter coils in a three-transmitter electromagnetic articulometer system designed by Perkell (see Perkell et al., 1992). Also shown are three receiver coils (transducers) attached to the tongue. Reproduced from Perkell et al. (1992) with permission.

In these systems, the axes of the transducer (i.e., receiver) coils are mounted perpendicular to the midsagittal plane and parallel to the transmitter axes. If the axes of the transmitter and receiver coils remain parallel, the strength of the magnetic field (and hence the voltage detected by the receiver coil) is inversely proportional to the cube of the intercoil distance (see Hixon, 1970). This relation makes it possible to determine the time-varying positions of the articulator fleshpoints. However, if there is any tilt in a receiver coil (which may occur due to articulator motion), the intercoil distance will be calculated as being larger than it really is. The three-transmitter systems currently in use include specialized hardware and software designed to detect and correct for tilt (Perkell et al., 1992; Schönle et al., 1987).

Like the microbeam system, this technique allows one to obtain large quantities of data on articulators inside (e.g., tongue and velum) and outside (e.g., lips) the vocal tract efficiently and, usually, without any discomfort to the subject. Current systems can track ten receiver coils (Carstens, Inc., personal communication, 1992; Perkell et al., 1992). With two coils used for reference (e.g., one on the bridge of the nose and one on the upper incisors), eight are available for articulator fleshpoints. If two are positioned on the lips (one on the upper and one on the lower lip), one on the lower incisors (for jaw tracking), and one on the velum, three are available for tracking the tongue. (Alternative arrangements of the coils are also possible.) In contrast to microbeam systems, magnetometers are commercially available and involve less cost (Perkell et al., 1992).

Electromagnetic articulometers are relatively new, and we are aware to date of only one study using this instrumentation to track velum movement. Katz, Machetanz, Orth, and Schönle (1990) examined movements of the velum, tongue, and lip in normal and apraxic subjects. The velar coil was attached with glue through the mouth to the inferior surface of the velum with the subject seated in a dental chair and tilting his head back.[7] (A coil can also be attached to the superior surface of the velum.) The data obtained provided movement trajectories that have the appearance of trajectories obtained with other movement transduction systems. Actual comparisons among data obtained with X-ray microbeam, optoelectronic transduction, and articulometry have yet to be undertaken, but it is clear that such comparisons are important.

Just as there have been rising concerns about the effects of electromagnetic fields from such sources as video display terminals (VDTs), electric blankets, and high-tension power lines, questions are being raised about electromagnetic articulometry. To date, however, the evidence suggests that these devices emit no more electromagnetic energy than such commonly used appliances as hair dryers and VDTs (Perkell et al., 1992).

3.2.3. VELOTRACE

The Velotrace is a mechanical device, developed by Horiguchi and Bell-Berti (1987), that tracks velum position over time. The device, inserted through the nose, consists of three major parts: a curved internal lever that rests on the nasal surface of the velum, an external lever that remains in full view outside of the nose, and a push rod (carried on a support rod) that connects the internal and external levers (Figure 12). Movements of the velum result in changes in the angle of the internal lever with respect to its fulcrum that are reflected in corresponding angular movements of the external lever; the levers are connected so that when the internal lever is raised (by a raising movement of the velum), the external lever moves toward the subject. Hence, measurement of movement of the external lever in the *x*-dimension provides information on the *y*-dimension of velum displacement. The external lever of the Velotrace is considerably longer than the internal lever, so the obtained displacements are larger than the actual displacements.

The Velotrace is inserted after the application of a topical anesthetic and decongestants to the nasal mucosa and posterior pharyngeal wall. The fulcrum of the internal lever is positioned in the nasal cavity at the end of the hard palate, with the tip of the internal lever resting on the nasal surface of the velum and the support rod on the floor of the nasal cavity. A special headband keeps the Velotrace stable in position, with clamps connecting the Velotrace to the headband.

An optoelectronic tracking system (Kay, Munhall, Vatikiotis-Bateson, & Kelso, 1985) is used to track the movements of light-emitting diodes (LEDs) mounted on the end of the external lever and on the fulcrum of the Velotrace (for reference). The positions of the LEDs in the sagittal plane are tracked by a position-sensitive detector whose output can be converted into pairs of (*x*, *y*)-coordinates for each LED. Using the optoelectronic tracking system, one can also attach LEDs to the lips and/or jaw (and the bridge of the nose, for reference) and thus collect data on several articulators simultaneously. Acoustic data are often collected along with the movement data as well (Figure 13).

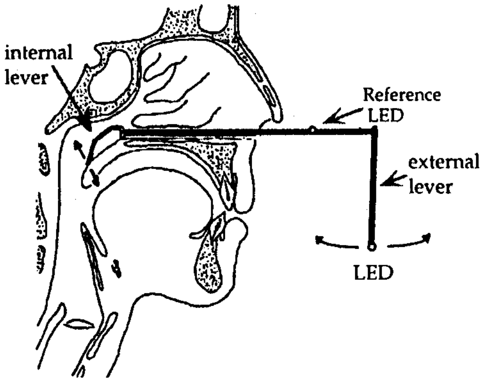

Figure 12. Schematic of the Velotrace. Adapted from Bell-Berti and Krakow (1991) with permission.

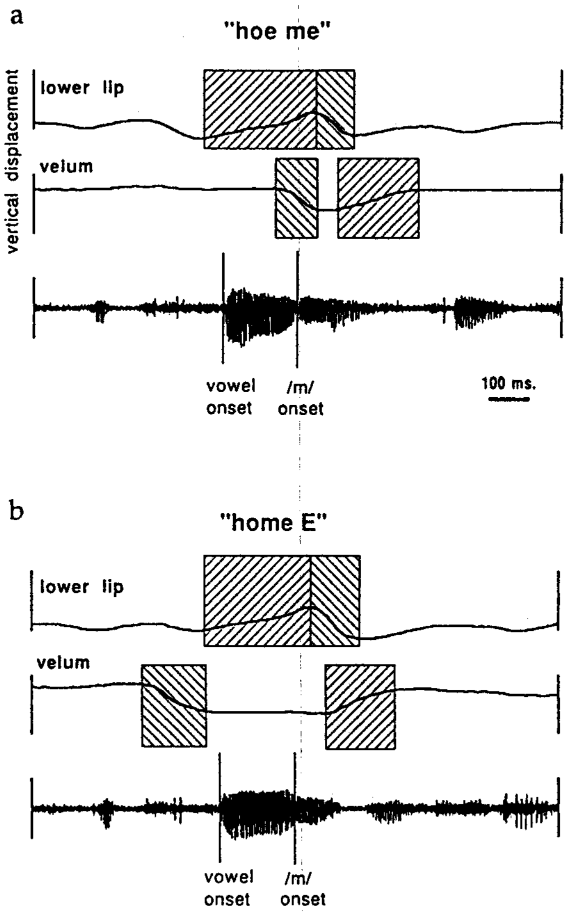

Figure 13. Vertical movements of the velum and lower lip obtained by Krakow (1989) in an experiment using the Velotrace and Selspot System. (a) shows the movements and corresponding acoustic waveform for *hoe me* and (b) shows the movements and acoustic waveform for *home E* from a study designed to examine the effects of word and syllable boundaries on velum–lip coordination. Movements of the articulators toward and away from the target positions for the nasal consonants are highlighted. (The data from this study are described in Krakow, this volume.)

In contrast to fiberoptic or cinefluorographic data, Velotrace signals have the advantage of not involving any frame-by-frame tracking. The device cannot, however, be calibrated, because as the velum moves, the precise point on its surface that is tracked by the internal level of the velum will change. Nonetheless, comparisons of Velotrace data with fiberoptic and cinefluorographic data on vertical velum displacement indicate considerable agreement (Horiguchi & Bell-Berti, 1987). Studies using the Velotrace include Bell-Berti (this volume), Bell-Berti & Krakow (1991), Horiguchi and Bell-Berti (1987), Krakow (1989, this volume), and Kollia (to appear).

3.2.4. STRAIN GAUGE

Strain gauges have been used in a large number of studies of lip (upper and lower) and jaw movements (see Abbs & Watkin, 1976; Hixon, 1972) and a few studies of velum movements (see Hixon, 1972; Moller, Martin, & Christiansen, 1971). The general principle of strain gauge transduction is the same, regardless of articulator: Paired strain gauges are mounted on one end of a flexible metal strip (one on the topside and one on the underside), and the other end of the strip is anchored to a stable support. The free end of the metal strip is deformed when pressed by a moving articulator, causing tension in the strain gauge at the top of the free end of the strip and compression of the strain gauge beneath. These effects are transduced as a change in the resistance of the strain gauge pair, functioning as two arms of a Wheatstone bridge, when coupled to two other resistors. (See Baken, 1987, for further discussion of the electronics of strain gauges.)

The devices used by Moller et al. (1971) and by Hixon (1972) to transduce velum movement have strain gauges attached to one end of a thin metal strip and a spring wire attached to the other end (Figure 14). The spring wire is designed to extend to the center line of the oral surface of the velum. Vertical velic movements displace the spring, which bends the metal strip, causing a change in resistance in the strain gauges. The entire device is fixed in position by attachment to an orthodontic band stabilized around an upper molar. A tongue guard is also used to prevent the tongue from putting pressure on the strain gauges from below. Although the development of such devices for tracking velic movements has been described in several papers (Christiansen & Moller, 1971; Hixon, 1970; Moller et al., 1971), their use has been quite limited, making it hard to render a judgment regarding their accuracy.

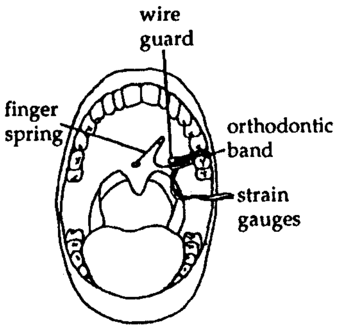

Figure 14. Schematic of a strain gauge designed to monitor velic movements. Adapted from Hixon (1972) with permission.

3.2.5. PHOTODETECTION

Photodetectors typically function to provide information about regions of the vocal tract that open and close, such as the glottis or the velopharyngeal port. These devices detect the relative amount of light that passes through the opening when a light source is situated above or below the opening with a light detector (a photocell) on the other side. Several such devices have been designed to determine relative extent of velopharyngeal port aperture (see, e.g., Dalston, 1982; Ohala, 1971).

The Nasograph (Ohala, 1971), for example, consists of a compressible polyethylene tube containing both the DC light source and the photodetector (Figure 15a). The tube, 4 mm in outside diameter, is inserted through the nostril, and the internal end is swallowed into the esophagus, which stabilizes the Nasograph with the light source situated below the velopharyngeal port (approximately at the level of the epiglottis or just above) and the light sensor above it (about at the junction between the hard palate and the velum). Time-varying information about velopharyngeal aperture is obtained in the form of voltage transduced from the amount of light detected coming through the port: with increased port opening, there is an increase in voltage. Although the device is not calibrated, it does provide useful information about relative port opening over time (Figure 15b; Clumeck, 1976; Ohala, 1971).

Another device that places a light source below and light detector above the velpharyngeal port was developed by Dalston (1982). This device is constructed by coupling a fiberoptic light source and an integrated photodetection assembly to a single coated plastic fiber, with an outside diameter of 0.75 mm. The two fibers are cemented together with the internal end (i.e., the end inserted into the subject's nose) of the light source extending 30 mm beyond the internal end of the photodetector pickup fiber. The internal ends of the light source and photodetector are in the approximate locations of the inserted light source and detector in Ohala's Nasograph.

Testing the device, Zimmerman, Dalston, Brown, Folkins, Linville, and Seaver (1987) noted that no discernible photodetector ouput was seen during the production of sustained phonemically oral vowels (produced by English speakers) in oral syllables, irrespective of vowel quality. Although this result suggested to them that there may be complete velopharyngeal closure for all oral vowels, radiographic and endoscopic techniques typically find shallow port opening for "oral" vowels,

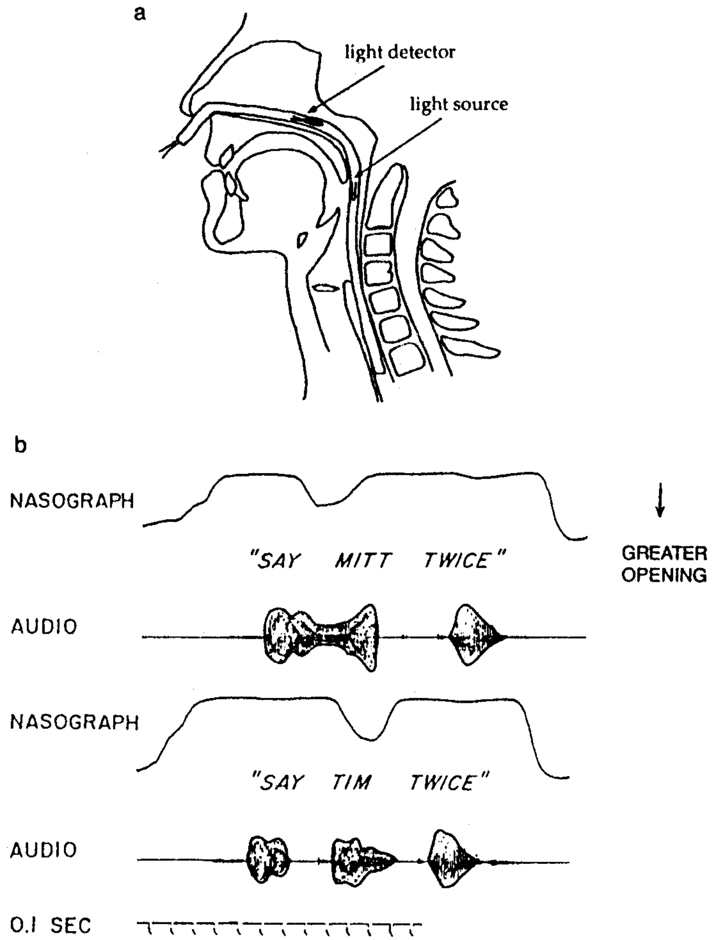

Figure 15. (a) Schematic of the Nasograph, a photoelectric device used to determine relative velopharyngeal aperture. (b) Nasograph data showing a difference in the magnitude of velopharyngeal aperture for the syllable initial nasal in *mitt* vs. the syllable-final nasal in *Tim*. Adapted from Ohala (1971) with permission.

with greater opening for low than for high or mid vowels (see Bell-Berti, this volume, and Figures 7 and 8). Examination of data obtained with the same device in another study (Dalston, 1982) as well as with the Nasograph (cf. Figure 15a) suggest that a certain minimum aperture threshold is required before light will pass or be detected coming through the open port. Still, differences in velopharyngeal port aperture due to factors such as the syllable position of a nasal consonant are evident in photodetector data (as in the Nasograph data shown in Figure 15b), along with differences between oral and nasal sounds, suggesting that after the threshold is reached, these devices can distinguish relatively small differences in the magnitude of port opening.

In contrast to the Nasograph and Dalston's device, the Velograph, a photodetector developed by Kuenzel (1977), was designed to be sensitive to changes in velum position rather than port opening. The Velograph consists of a thin, flexible probe which is inserted through the nostril and positioned so that light is emitted from its internal end onto the nasal surface of the velum. The internal end is positioned approximately where Ohala's light detector is positioned (see Figure 15a). With the Velograph, the light is reflected off of the surface of the velum, as a function of velic height, and the reflection is picked up by a miniature photocell also contained in the tip of the probe. In contrast to the photodetectors of Dalston and Ohala, no part of the device passes through the velopharyngeal port, and the Velograph *does* produce data showing vowel height effects for oral vowels (Figure 16). Because the velum continues to rise even after port closure has been achieved, devices designed to detect velic height have generally been more sensitive to differences among oral segments than have devices designed to detect velic port opening.

There are several serious issues to be considered, however, in evaluating the Velograph. First, the amount of velic surface area from which the light will be reflected is a function not just of velum height, but also of pharyngeal wall movement. Second, the collection of mucus in the region, which commmonly appears and has a highly reflective surface, may result in the acquisition of spurious data on velum position from the Velograph. This suggests that the use of a photo-

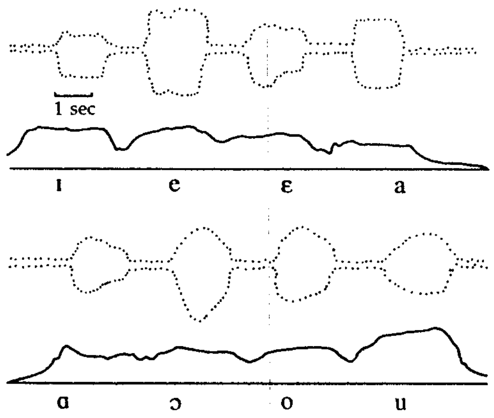

Figure 16. Data on velum height for different vowels obtained with the Velograph. Adapted from Kuenzel (1977) with permission.

detector to measure aperture is probably more appropriate than its use to track velum height.

In general, photodetectors are safe, and with proper insertion and placement procedures there is no physical risk and minimal, if any, discomfort. Thus, large amounts of data can be obtained. These devices are also inexpensive and relatively easy to position. (In some of these studies, the subjects positioned the device themselves.)

4. MOVEMENT EFFECTS: AERODYNAMICS AND ACOUSTICS

Previous sections of this article have discussed techniques for studying muscle activity and articulatory movements that affect velopharyngeal port opening. In this section, we discuss techniques for studying some of the aerodynamic and acoustic consequences of articulator movements. We focus on nasal and oral airflow, vibration of air in the nasal cavities (nasal resonance), overall spectral characteristics of speech sounds, and what these kinds of data can tell us about velopharyngeal function.

As discussed in previous sections, some procedures for collecting data on velopharyngeal port aperture are invasive and may be uncomfortable, and a few may have some risk associated with their use. In contrast, procedures for collecting aerodynamic and/or acoustic data are usually noninvasive, cause no discomfort, and are completely safe for the subject. This in turn means that, with aerodynamic or acoustic techniques, it is more feasible to conduct large-scale studies, including more subjects (including children) and possibly more languages in the data sample. In addition, the equipment needed to record aerodynamic and acoustic data tends to be more portable than that used in physiological studies of speech. This gives the experimenter further flexibility in experimental design, since subjects need not be limited to those who can travel to the laboratory. For the same reason, portability is also an advantage in clinical applications.

On the other hand, caution must be exercised when planning and carrying out experiments which investigate these more indirect indicators of velopharyngeal function. For example, we know that some articulatory changes are not reflected straightforwardly in aerodynamic and acoustic measures. Beyond a certain degree of velopharyngeal opening, changes in velum position have a nonlinear effect on the amount of nasal airflow and a minimal effect on the spectral properties of the speech produced. Furthermore, there can be changes in aerodynamic or acoustic data that do not reflect changes in velopharyngeal opening and nasal coupling per se. These possible confounding factors can be eliminated or, at least, minimized through the use of appropriate experimental controls, as is discussed in the presentation of individual techniques.

In the sections that follow we discuss different kinds of aerodynamic and acoustic data individually, but it should be noted that it is often profitable to collect a combination of these types of data. If collected and interpreted carefully, aerodynamic and acoustic data can provide valuable information about velopharyngeal control and nasal coupling.

4.1. Airflow

When the velopharyngeal port is open during speech, under normal conditions air flows through the nose. In speakers without velopharyngeal abnormalities, the volume of air that flows through the nasal cavities is related to the size of the velopharyngeal opening. Thus, nasal airflow is an indicator of velopharyngeal function in normal subjects and may be used to gain insight into the linguistic and physiological factors affecting normal velopharyngeal function (see, e.g., Benguerel, 1974; Cohn 1990, this volume; Huffman, 1989, this volume). Moreover, speakers with velopharyngeal abnormalities or neuromuscular pathologies may show patterns of nasal airflow that are different from those observed on comparable speech items for normal subjects. For this reason, airflow data are often used in the diagnosis and evaluation of conditions involving both excessive nasal resonance (hypernasality) and insufficient nasal resonance (hyponasality), as discussed in Warren, Dalston, and Mayo (this volume).

Airflow data relevant to studies of nasalization are obtained by channeling the air into a pressure measurement device that is used to transduce the airflow into a time-varying electrical signal. Typically, the airflow measurement device is built into a mask which is placed over the nose or the nose and mouth. In the latter case, oral and nasal airflow are kept separate by a divider built into the mask; such a mask is called a *split mask*. Another commonly used device has a pressure transducer attached to a tube inserted in one nostril; a cork or a rubber plug through which the tube passes is used to seal the nostril and secure the tube. (See Baken, 1987, Chap. 8, for an excellent discussion of the technical details of these devices.) One disadvantage of using airflow masks is that they sometimes interfere with jaw lowering and lip movement (Lubker & Moll, 1965), but there are also a number of important advantages: in addition to their safety and noninvasiveness, these devices also produce data that can be calibrated, permitting measurements of speech behavior that are useful for both clinical and linguistic studies.

When using nasal airflow data to make inferences about velopharyngeal aperture, one must control for factors that may cause changes in nasal airflow that do not reflect changes in velopharyngeal opening and nasal coupling. The two major confounding factors affecting the interpretation of nasal airflow data are changes in oral resistance to airflow and changes in overall airflow.

The amount of air flowing through the nose depends not only on velopharyngeal port size, but also on whether, and how much, air is flowing through the mouth. Assuming a constant airflow source and a constant velopharyngeal opening of moderate size, increased resistance to airflow through the oral cavity (as a result of partial or complete oral constriction) will increase the amount of airflow through the nasal cavities. As a result, when the amount of nasal airflow changes over the course of a segment, one must consider whether this may be due to changes in oral constriction. Acoustic data (e.g., spectrograms) or physiological data on the oral articulators can provide this information and thus aid in the interpretation of nasal airflow. There is no simple way, however, to factor out the effect of changes in oral constriction. Another consequence of the influence of oral constriction on nasal airflow is that one cannot compare absolute nasal airflow levels for sounds differing in degree of oral constriction, such as [i] and [a]. However, it is appropriate to compare absolute nasal airflow levels on one segment in minimally contrastive oral and nasal contexts. For example, airflow during the vowel of *bid* could be compared to that during *mid* or *bin* in English. In fact, data on velopharyngeal function and nasal resonance are best interpreted in relative terms, comparing nasal items to matched oral control items whenever possible, as discussed in Section 1. The issues of data interpretation raised here must be considered at the time of experimental design, so that the necessary supplemental data (e.g., audio or articulatory recordings) are collected and the speech materials include the appropriate phonetic contexts for the types of comparison that are needed.

Another important factor which must be considered in nasal airflow studies is that nasal airflow will be affected by the overall amount of airflow. Overall airflow will change with variations in glottal aperture—such as the increase in aperture that takes place in the production of fricatives and voiceless and/or aspirated sounds—and changes in subglottal pressure—such as the increased pressure that often accompanies lexical stress, or the decreased pressure that occurs near the end of an utterance.

One way to control for the effects of possible changes in overall airflow is to determine the proportion of overall flow which is passing through the nose. This requires recording oral as well as nasal airflow (e.g., using the split mask) and then computing the quantity *nasal airflow/(nasal airflow + oral airflow)*. This proportion will be affected by changes in velopharyngeal port size and/or oral constriction but will be unaffected by changes in overall flow. Of course, one might choose not to make an oral airflow recording, either because of instrumental limitations (e.g., the appropriate device is not available) or because a good-quality audio recording is desired. In such cases, nasal airflow would be recorded with a device that does not cover the mouth, such as a tube inserted into the nostril, a technique discussed earlier. If oral airflow is not recorded, the speech materials used must be designed to avoid sounds which are likely to involve large changes in overall airflow—such as [h] and aspirated stops—and care should be taken to make comparisons only between items similar in stress and position in the utterance, because overall flow can also vary with these factors.

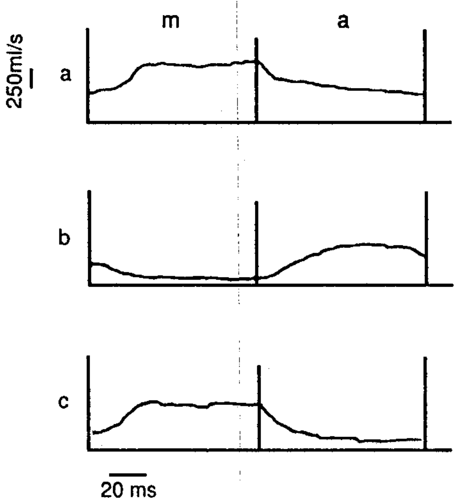

Figure 17. Examples of (a) raw nasal airflow, (b) raw oral airflow, and (c) proportional nasal airflow for one repetition of /ma/, as produced by a speaker of Akan. The vertical lines mark acoustic segment boundaries, based on spectrograms and waveform displays of a simultaneous audio signal.

Figure 17 illustrates airflow data for a speaker of Akan (a language of Ghana). The data represent one repetition of the syllable /ma/, from /oma/ 's/he wants'. Figures 17a and 17b, respectively, show the raw nasal and oral airflow traces for this item, and Figure 17c shows the output of a computer program that computes the proportional nasal airflow signal from the nasal and oral airflow signals. [The slight offset from zero flow—clearly visible on the oral airflow trace during the /m/—is due to the rms (root mean square) algorithm used to smooth the raw airflow signals.] Because the *y*-axis of the curve in the proportional nasal airflow curve represents percentages, values along this dimension cannot be compared directly with values along the *y*-axes in the raw nasal and oral airflow signals, which represent airflow in ml/sec.[8] However, the time scale is the same in all three curves, and changes in the shape of the proportional nasal airflow curve can be taken as meaningful. We will sketch briefly some of the issues involved in interpreting proportional nasal airflow data.

First, it is important to note that the proportional nasal airflow curve is only informative for sounds that do not involve complete oral closure. For sounds produced with complete oral closure and some amount of nasal airflow (such as [m]), the proportion of nasal to total flow is always 100%, regardless of the absolute amount of nasal airflow, because in this case all airflow is through the nose. Thus, changes in nasal airflow during such consonants will not be evident in the proportional airflow signal.

Turning to the /a/ of /ma/, it can be seen that the proportional nasal airflow curve shows an initial drop after the release of the /m/, followed by a further drop which gradually levels off about three-fifths of the way into the vowel, after which the trace is flat for the remainder of the vowel. When proportional nasal airflow is changing, as in the first part of the /a/ in Figure 17, this may be due to changes in velopharyngeal port opening, changes in oral constriction, or both, but we know that it is not due to changes in overall flow (because the latter changes have been factored out). Because oral aperture increases at the release of the /m/, we know that a change in oral constriction is at least one factor contributing to the drop in the proportional nasal airflow curve in Figure 17c. A decrease in velopharyngeal port area may also be contributing to the drop in proportional nasal airflow. This hypothesis could be investigated using supplementary data, such as by analyzing a simultaneous audio recording for spectral evidence of decreasing velopharyngeal port area (e.g., change in frequency of the nasal formant and/or antiformant; see Section 4.3.2) or by using aerodynamic modelling of pressure and airflow recordings to estimate port area (e.g., Warren & Devereaux, 1966; Warren & DuBois, 1964).

When proportional nasal airflow is not changing, as at the end of the /a/ in Figure 17c, this means that velopharyngeal port area and oral aperture are behaving similarly. That is, either both remain unchanged, or they are changing together (both increasing or both decreasing), at approximately the same rate, with the result that the proportion of air flowing through the mouth and the nose does not change. Again, choosing between the alternative interpretations requires additional data. Either acoustic or physiological data on the oral articulators could indicate whether there was a change in oral constriction during the last part of the vowel. If there were evidence of such a change, this would favor the interpretation that both the mouth and the velopharyngeal port are closing late in the vowel. If such evidence were lacking, this would favor the interpretation that oral aperture and velopharyngeal port size remain basically unchanged during this part of the vowel. It should be noted that, while we have discussed airflow patterns of one token in some detail here, when carrying out an experiment one would normally have different utterances and multiple tokens of each utterance to compare, which would aid the process of data interpretation by revealing which properties of an utterance are the result of context, which are idiosyncratic, etc.

4.2. Sound Pressure

When there is a passage into the nasal cavities, under normal aerodynamic conditions (e.g., sufficient air pressure, only moderate nasal resistance to airflow) the air in the nasal cavities resonates.[9] The resultant air pressure variations can be detected with a number of methods. One approach uses a microphone to detect sound pressure in the vicinity of the nostrils, since nasal sound pressure increases with nasal resonance. Another approach uses a contact microphone (accelerometer) to detect nasal resonance by recording the vibrations of the soft tissues of the

nose; these vibrations are generally stronger when there is nasal resonance. With either of these methods, the result is a time-varying signal.

As with airflow measures, certain experimental controls are recommended to aid in the interpretation of data on nasal vibration. For example, the experimenter must take into account the fact that there can be some nasal vibration even in the absence of velopharyngeal opening and nasal coupling, because sound vibrations can be conducted from the oral cavity to the tissues of the nose. There is some evidence that this vibration is stronger for sounds produced with greater oral constriction (e.g., high vowels). As with airflow data, then, it is important to consider the behavior of sounds in a linguistically nonnasal context as a reference against which to evaluate the data on sounds in a nasal context, and care must be taken in making comparisons between utterances which involve sounds differing in degree of oral constriction. In addition, it is important to distinguish changes in nasal resonance from general changes in the speech system output. How this is done varies slightly with the device being used, as detailed below in the individual sections on sound pressure microphones and accelerometers.

4.2.1. SOUND PRESSURE MICROPHONES

Nasal sound pressure variations can be detected by a microphone positioned inside a nostril or just outside a nostril. When the microphone is placed inside the nostril, it is often secured in a stable position by threading the microphone wire through a plug that is then inserted into the nostril. Another method of stabilizing the nasal microphone inside the nose is to attach the microphone to a stationary object and position the subject's head so that the microphone sits inside the nostril.

As noted above, care must be taken to distinguish changes in nasal vibration from changes in overall vocal output. For example, if a subject's speech becomes louder, there will be an increase in the amplitude of nasal vibrations, but this cannot be taken to mean that there has been an increase in velopharyngeal opening and nasal coupling. Information about overall output can be obtained by recording sound pressure variations at the mouth simultaneously with the nasal recordings. The oral signal can then be used in interpretation of the nasal signal. One way to do this is to express the amplitudes of the nasal and oral signals as a mathematical relation such as a difference or a ratio, as described above for airflow data.

When the microphone used for recording nasal vibration is placed outside the nose, one must be particularly careful to ensure that the nasal signal is isolated from the oral signal emitted from the mouth. A barrier may be constructed for this purpose, or one can use a device that already has a barrier built in, such as the Nasometer (an adaptation of the TONAR system; Fletcher, 1970). As shown in Figure 18a, the Nasometer has two microphones, one mounted on the upper and one on the lower surfaces of a dividing plate. The plate is placed at the level of the upper lip and is stabilized by attaching it to a head brace. In the Nasometer sys-

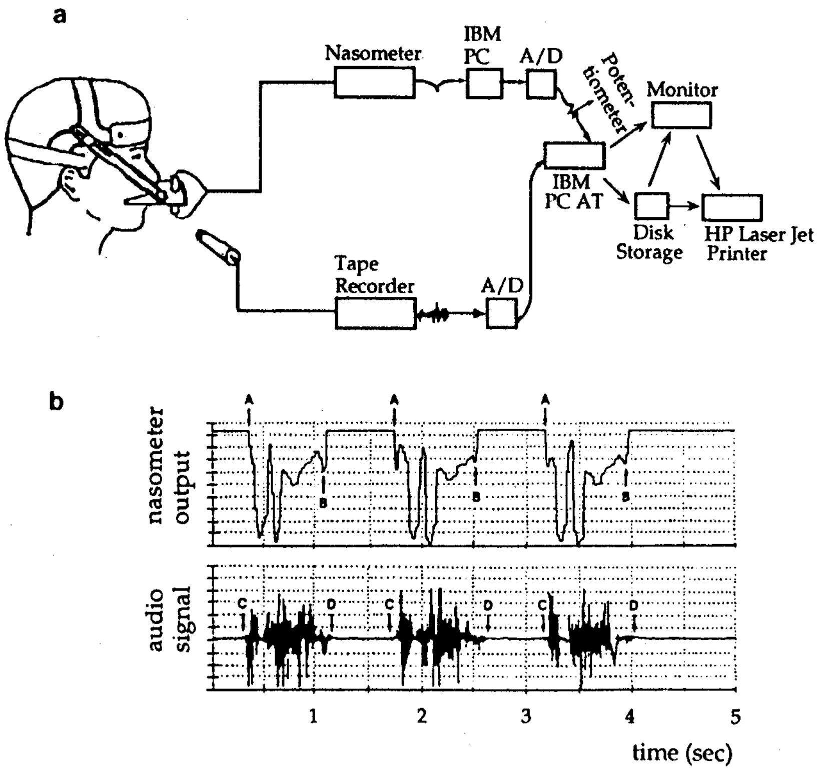

Figure 18. (a), Schematic of an experimental setup including the Nasometer. Oral and nasal microphones are mounted on opposite sides of a dividing plate which is held in place above the upper lip by a head brace. (b), Nasometer output and a simultaneously recorded audio signal for three repetitions of the phrase *Come to my house* as produced by a normal adult. Nasometer output has been inverted, so increased nasalance appears as a downward excursion in the curve. A and B mark the onset and offset of each repetition in the nasometer signal. C and D mark the same events in the acoustic signal. Reproduced from Seaver and Dalston (1990) with permission.

tem, the relation of nasal and oral signals is computed as a ratio of nasal to nasal-plus-oral energy, multiplied by a constant (100). The resulting quantity is termed *nasalance* and has been found to be useful in clinical diagnosis of velopharyngeal impairment (e.g., Dalston, 1989; Dalston, Warren & Dalston, 1991; Warren, Dalston & Mayo, this volume).

The safety, noninvasiveness, and ease of use of the Nasometer system are also significant factors in its increasing application in clinical settings. Figure 18b illustrates Nasometer output and a simultaneously recorded audio signal for three repetitions of the phrase *Come to my house.* Nasometer output has been inverted, so increased nasalance appears as a downward excursion in the curve. For each repetition of the phrase, one can see a sharp increase in nasalance for the nasal consonants in *come* and *my.* The Nasometer appears to perform well in detecting the fairly strong nasal resonance associated with hypernasal speech or that found for nasal consonants in normal speech. However, the Nasometer system appears also to respond to acoustic events other than nasal resonance. For example, Seaver and Dalston (1990) note that they sometimes obtained a positive Nasometer signal on oral sounds, particularly voiceless consonants lacking strong frication. Further study is needed to determine the reason for this effect and how it might be eliminated.

One technical disadvantage of using microphones to measure nasal sound pressure is that they are affected by airflow changes as well. As a result, nasal airflow passing over the microphone is confounded with sound pressure variations. This affects how the signal is interpreted and limits its usefulness. This problem does not occur with accelerometers, which differ from standard microphones in that they are designed specifically to follow and transduce vibrations of the surfaces to which they are attached, rather than directly registering air pressure variations.

4.2.2. ACCELEROMETERS

When an accelerometer is used to measure nasal vibrations, it is placed on the exterior surface of one nostril and held in place with an adhesive, such as double-sided tape. The signal is then processed in the same way as a sound pressure microphone signal. The accelerometer's insensitivity to airborne sound waves gives it a number of advantages over sound pressure microphones. First, it is relatively unaffected by oral airflow and sound emission, thus more effectively isolating the nasal signal. In addition, the accelerometer signal is not affected by nasal airflow. Stevens, Kallikow, and Willemain (1975) and Stevens, Nickerson, Boothroyd, and Rollins (1976) report successful use of accelerometer signals as a source of visual feedback on nasality for deaf children.

Despite these advantages, the accelerometer has its own drawbacks which must be considered when planning experiments which use this device. First, the accelerometer is quite sensitive to differences in the mechanical properties of the tissues to which it is attached (Baken, 1987). This means that absolute values of signal amplitude cannot be compared across speakers, or for that matter, across sessions with the same speaker, because accelerometer placement will inevitably vary between experimental sessions. Second, accelerometers show some sensitivity to oral tract resonance—accelerometer signal level has been observed to increase with a lowering of the first formant (F1; Larson & Hamlet, 1987; Stevens et al., 1976), as would occur with increased oral constriction. This means that one should avoid comparing absolute levels of accelerometric signals for sounds that have different F1 values. It also suggests that one must be careful in interpreting accelerometer signals in periods of speech when F1 is changing rapidly—such as when degree of oral constriction changes during the transition between a stop consonant and a vowel (see Larson & Hamlet, 1987)—because in such a case a change in the accelerometric signal might not indicate a change in velopharyngeal opening or nasal resonance. Supplementary data such as an audio recording can aid in the interpretation of the accelerometer signal by providing additional information about changes in oral constriction.

Like the air pressure microphone, the accelerometer measuring nasal vibration should be used simultaneously with a device that can provide information about overall vocal output, to help in identifying changes in nasal vibration which are exclusively due to changes in velopharyngeal opening. With an accelerometer system, this is usually done by making an additional accelerometer recording on the neck at the level of the glottis (although one might also use a regular acoustic recording made with a microphone near the lips, as in Larson & Hamlet, 1987). The glottal signal can then be compared to the nasal signal, or it can be used to compute the relation of nasal and glottal (oral) signal amplitude.

Horii (1980) has proposed a strategy for correlating the nasal and oral accelerometer signals that provides a control for changes in vocal output and also, he claims, allows for meaningful cross-speaker and cross-session comparisons. In computing this HONC (Horii Oral Nasal Coupling) index, the rms nasal signal amplitude is divided by a quantity that is the rms oral signal amplitude multiplied by a constant. The constant is determined for each experimental session so that the HONC index is 1 for maximal nasalization produced by the subject during the session's calibration procedure. Since the total scale used (0 to 1) is the same for every recording session, Horii suggests that results from different sessions should be comparable, whether they involve the same speaker or different speakers.

Figure 19 shows nasal and oral accelerometer signals and the corresponding HONC index for the phrase *acoustical nasal coupling index* obtained by Horii (1980); in this case, the HONC scale is displayed in decibels. In the lower curve, representing the HONC values, the largest positive excursions correspond to the three nasal consonants in the phrase; there are much smaller excursions for oral vowels in oral consonant contexts, such as the [u] of *acoustical* and the second vowel of *nasal*. Note also that the HONC index is not computed for voiceless sounds. This is because when there is no voicing, dividing the nasal signal by the oral signal (to obtain the HONC index) would mean dividing the nasal signal by zero. Horii (1983) and Redenbaugh and Reich (1985) report that HONC values are well correlated with perceived nasality.

It should be pointed out, however, that changes in the HONC index do not always indicate changes in nasal resonance because, being a nasal/oral ratio, the HONC index can be affected by changes in the oral signal as well. Such an effect is reported by Matthies, Perkell, and Williams (1991), who discuss longitudinal data on HONC values for cochlear implant users. After a year of using a cochlear implant, two subjects showed an increase in HONC values accompanied by an improvement in overall perceived voice quality. If increased HONC values indicate increased nasality, it is surprising that the subjects' overall voice quality was judged to have improved. Spectral analysis of the accelerometer signals suggested that, rather than indicating increased nasality, the higher HONC values were due to a reduction in the amplitude of the throat accelerometer signal, caused by a change toward more breathy or less pressed phonation. Nasal airflow data supported the conclusion that there had not been an increase in nasality: nasal airflow did not increase with the increase in the HONC index.

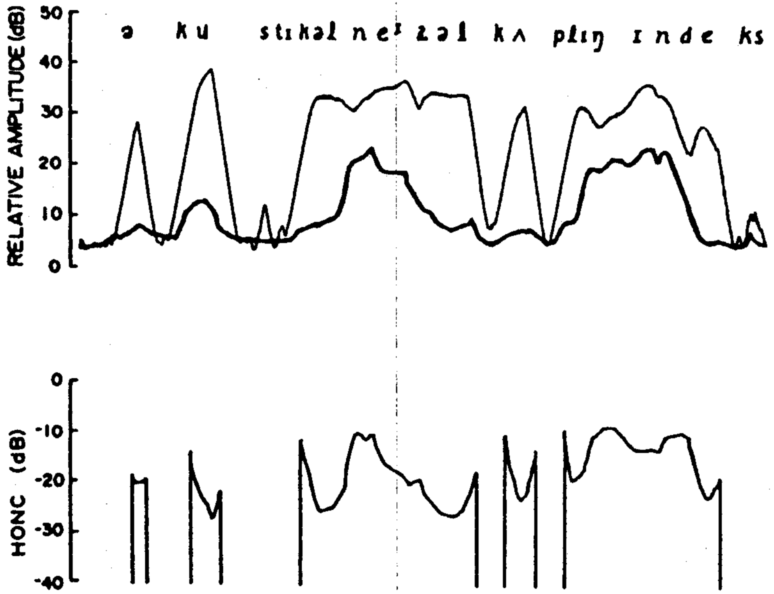

Figure 19. Nasal and oral accelerometer signals and the corresponding HONC index in dB for the phrase *acoustical nasal coupling index*. Nasal (*heavy line*) and oral (*thin line*) accelerometer signals are superimposed in the upper trace. Adapted from Horii (1980) with permission.

4.3. Spectral Analysis

When there is velopharyngeal opening and air flows through the nose during speech, the nasal cavities form part of the resonating chamber, contributing to the spectral properties of the sound radiated from the vocal tract. Given the ease and noninvasiveness of making audio recordings, spectral analysis is a very attractive tool for studying nasalization. However, for a variety of reasons, it has been difficult to arrive at a single reliable spectral measure of nasalization. One reason is that when the nasal tract is coupled to the oral tract, the resonances of the two systems interact in acoustically complex ways (see, e.g., the articles by Maeda and Ohala & Ohala, this volume). A further complication is the fact that there are substantial differences across speakers in the shape of the nasal tract, meaning that results of spectral analysis for one speaker may not predict well what will be observed for other speakers. Nonetheless, with an understanding of the basic principles involved, spectral analysis of the acoustic signal can provide useful information about the presence and amount of nasalization. Here we will concentrate on the spectral consequences of nasalization of vowels and how this kind of data can provide insight into velopharyngeal aperture and function. For discussion of spectral properties of nasal consonants, the reader is referred to Kurowski and Blumstein (this volume) and additional references provided there.

4.3.1. Spectral Characteristics of Vowel Nasalization

Many studies of the acoustics of nasalization have had the goal of identifying a single spectral property which is characteristic of vowel nasalization, independent of vowel identity. Thus the approach often taken in such studies has been to observe the spectral properties of individual nasalized vowels and then to try to identify the spectral characteristics that are common to all of them.

The spectral properties of nasalized vowels are usually described in terms of how they differ from those of oral vowels. In this view, the primary acoustic effect of coupling the nasal tract to the oral tract is the presence of additional resonance and antiresonance pairs. How these affect the spectrum of the oral vowel depends primarily on their frequency and how they are distributed in frequency relative to the resonances of the oral tract. Thus the effects of nasal coupling vary for different vowel qualities (for more detailed discussion, see Stevens, Fant, and Hawkins, 1987, and Maeda, this volume). In general, there are often changes in the frequency and bandwidth of the oral vowel's formants, especially F1, and there are additional formants and antiformants in the spectrum. It should be noted, however, that coupling the nasal tract to the oral tract results in a single acoustically complex system, rather than the combination of separate oral and nasal systems that is implied by this approach (see Curtis, 1970, for further discussion). However, this way of characterizing nasal vowel spectra, though an oversimplification, is convenient and for many purposes, sufficient. It is the approach taken in much of the work on the acoustics of nasalization.

The acoustics of vowel nasalization have been investigated using both speech analysis and vocal tract models. A few studies (e.g., Fujimura & Lindqvist, 1971; Lindqvist-Gauffin & Sundberg, 1976) have analyzed vocal tract acoustics in the absence of speech, by exciting the air in the vocal tract with electrically generated tones at various frequencies. An early study which proved to be very influential was that reported by House and Stevens (1956). Using an electrical analog of the vocal tract, they modeled nasal coupling on vowels and found substantial effects on F1, which exhibited reduced amplitude, increased bandwidth, and an upward shift in frequency. These effects are due primarily to the presence of an additional resonance–antiresonance pair near F1. The precise location of resonances and antiresonances, and the degree of formant shift, varied with vowel quality. House and Stevens also found evidence of additional resonances and antiresonances at higher frequencies, and they observed that overall vowel amplitude was reduced with nasal coupling. This latter effect is due in part to the damping properties of the nasal tract, which were simulated explicitly in the model. House and Stevens concluded that nasal coupling broadened the spectral peaks and generally flattened the vowel spectrum.

In another influential early study, Hattori, Yamamoto, and Fujimura (1958) examined spectral consequences of varying degrees of nasalization on vowels produced by two speakers of Japanese. They too found spectral changes in the region of the first formant, along with general spectral flattening. The principal characteristics of nasalization they observed on vowels were a "dull" resonance around 250 Hz, an anti-resonance at about 500 Hz, and additional weak components filling the valleys between formants.

Subsequent studies have provided additional data on a number of aspects of the acoustics of nasalization, including the spectral effects of nasalizing different vowels (e.g., Hawkins & Stevens, 1985; Maeda, this volume; Stevens et al., 1987) and the contribution of individual anatomical differences to characteristics of the spectrum (Dickson, 1962; Lindqvist-Gauffin & Sundberg, 1976). However, these results have not changed the overall picture provided by the early studies. The spectral variations due to vowel identity and speaker characteristics are numerous enough that a single description of the spectral properties of nasalized vowels as a whole must necessarily be worded in very general terms, along the lines of House and Stevens's conclusion that nasalization results in a general flattening of the vowel spectrum.

The spectra in Figures 20 and 21 illustrate some of the effects which occur with nasalization of the vowels [i] and [ɪ]. Figure 20a is a spectrum from the middle of the vowel [i] of the English *bead;* and 20b is a spectrum from a similar point in the vowel of the word *bean.* A comparison of these two spectra shows that the vowel of *bean* has more spectral energy in the valley between F1 and F2, centered between 500 Hz and 1000 Hz. This suggests that, in the /i/ of *bean,* the velopharyngeal port is open, coupling the nasal tract to the oral tract and resulting in an additional formant in this frequency range. The spectrum in 20c, representing a timepoint 20 msec after that in Fig. 20b, also shows this additional spectral energy between 500 and 1000 Hz. The latter two spectra illustrate how an antiresonance introduced by nasal coupling can affect the spectrum. In the later spectrum (c), harmonic amplitudes have decreased in the vicinity of 1500 Hz, while they have increased in the vicinity of 700 Hz. This suggests that there is an antiformant in the spectrum that has moved up in frequency during the 20 msec between the two

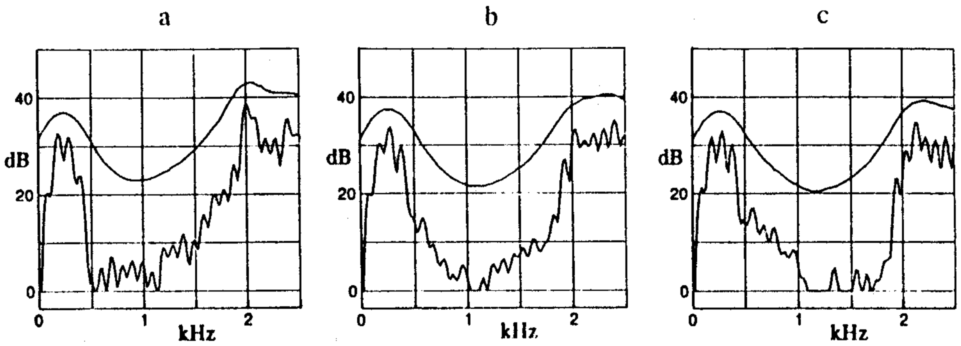

Figure 20. DFT spectra from (a) the middle of the vowel /i/ in the English word *bead;* (b) the middle of /i/ in *bean;* and (c) a point of 20 msec later in *bean.*

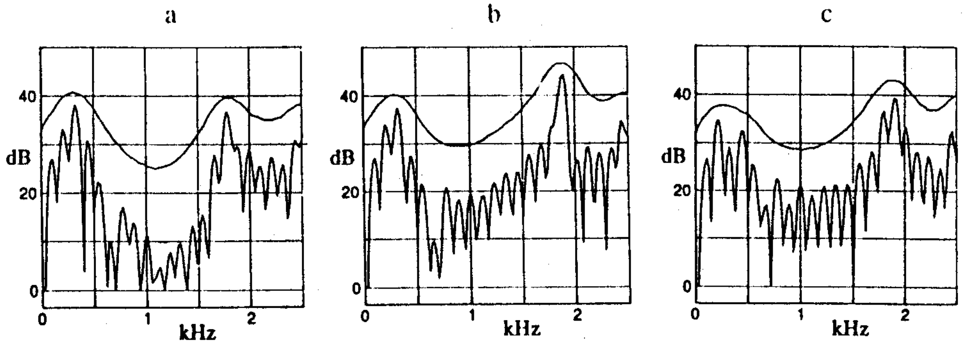

Figure 21. DFT spectra from (a) the middle of the vowel /ɪ/ in the English word *bid;* (b) the middle of /ɪ/ in *bin;* (c) a point 20 msec later in *bin.*

spectra. At the point in time represented in (b), the antiresonance is close to the added nasal resonance, near 700 Hz, reducing harmonic amplitudes in this frequency region; 20 msec later, the antiresonance has shifted up in frequency, reducing harmonic amplitudes around 1000 Hz while allowing harmonics around 700 Hz to rise by as much as 5 dB.

Figure 21 shows the corresponding spectra for the vowel [ɪ]. The spectrum in (a) is from the middle of the vowel of the English word *bid;* (b) is from the middle of the vowel in *bin,* and (c) is from 20 msec later in *bin.* There is little difference between (b) and (c), which suggests that, in the middle of *bin* (b), there is little or no velopharyngeal opening in anticipation of the nasal consonant. However, comparing the spectra in (b) and (c), we do see some evidence of velopharyngeal opening in the later spectrum. The spectrum in (c) shows higher harmonic amplitudes in the region from 400 to 700 Hz, suggesting that there is an additional resonance in this frequency region. F1 also appears to have shifted down slightly over the 20 msec between the timepoints represented in (b) and (c), since the second harmonic has increased in amplitude and the third harmonic has decreased in amplitude. The general effect that these changes have on the spectrum is that there is a broader prominence in the F1 region in (c) (i.e., later in the vowel), an effect typical of increased nasalization.

Because of the many factors affecting the spectra of nasalized vowels, attempts to quantify spectral properties of nasalized vowels have usually focused on characterizing the pattern of overall energy distribution or spectral slope that results from the extra resonances and antiresonances that come with nasal coupling. These measures are discussed in the next section.

4.3.2. SPECTRAL MEASURES OF VOWEL NASALIZATION

A spectral measure for clinical and linguistic study of nasalization should provide information not only about the presence or absence of nasalization, but also about changes in nasalization over time. However, most studies of the acoustics of vowel nasalization have generally focused on static, rather than dynamic, aspects of the signal. This is probably because attempting to find a single spectral property for nasalized vowels requires factoring out variation, including temporal variation. Nonetheless, a dynamic representation of nasalization can be derived using an essentially static spectral measure, by performing spectral analyses at regular intervals throughout the segment(s) of interest and plotting the results. So, for example, in the case of a vowel–nasal consonant sequence, nasalization of the vowel in anticipation of the consonant could be detected as an increase in degree of nasalization over successive analysis points during the vowel. However, such an analysis is only successful if used with a spectral measure that is relatively sensitive to differences in degree of nasalization. In this section we describe a few approaches to quantification of the spectral properties of nasalized vowels, and we evaluate their usefulness for deriving information about the presence and degree of nasalization and changes in nasalization over time.

Glass and Zue (1985) combined a variety of spectral measures into a nasalization detection algorithm for English vowels. In accordance with the general literature on the acoustics of nasalization, these measures focused on the low-frequency region of the spectrum, quantifying properties such as the presence of an extra resonance near the first resonance of the vowel and the relative magnitudes of these two resonances (see also Glass, 1984). To evaluate the effectiveness of their nasalization detection algorithm, Glass and Zue first divided English vowel tokens into "nasal" and "oral" categories, where vowels in the context of a nasal consonant were classified as "nasal", and all other vowels were classified as "oral." The detection algorithm was then tested on vowel tokens from six speakers (three male, three female), by first training on the vowels of five speakers and then being tested on vowels of the sixth. (This procedure was repeated a total of six times so that each speaker was used once as the final test case.) The algorithm classified vowels correctly an average of 74% of the time, with more cases of "oral" vowels being misclassified as "nasal" than vice versa. As Glass and Zue point out, these errors are not necessarily an indication that their measures fail in identifying nasalization; rather, they may be due in part to the fact that their definitions of "nasal" and "oral" categories ignore the fact that some speakers nasalize oral vowels (particularly the low vowels) even in oral phonetic contexts. Glass and Zue's measures could be identifying such cases as "nasal," which would mean that the measures produce the correct acoustic evaluation of the presence of nasalization, even though the resultant classification does not match Glass and Zue's a priori categorization based on phonetic context.

While Glass and Zue's algorithm may actually identify nasalized vowels fairly successfully, it is not a practical option for preliminary studies of nasalization in a general experimental setting. For one thing, their algorithm involves a complex of several spectral measures, which are evaluated in a computationally intense decision-making procedure. Furthermore, the algorithm is not designed to detect differences in degree of nasalization that can provide information about the time course of changes in velopharyngeal port opening and nasal coupling.[10] We turn next to several measures of vowel nasalization that are somewhat simpler to implement; however, they have been tested against more restricted datasets than Glass and Zue's measure.

Reetz (1991) describes a technique that uses linear prediction coding (LPC) analysis to determine the presence of a particular resonance that is indicative of nasalization, in this case a resonance below 400 Hz. It has been suggested (e.g., Lindqvist-Gauffin & Sundberg, 1976) that some low-frequency formants and antiformants in the spectrum are due to resonances of the frontal and maxillary sinuses. Reetz's metric may be detecting one of these low-frequency resonances, which are constant in frequency for a particular speaker. Reetz's algorithm involves performing a high-order LPC analysis, in 12.8-msec steps, through the voiced portion of a vowel. The LPC analysis must be high-order (computing a large number of formants in the designated frequency range) to make it possible to detect a nasal resonance which is close to F1 in frequency. The lower order LPC analysis often used in studies of vowel properties would normally find only one formant in the F1 frequency region; if an additional formant were present close to F1, it would be taken as part of F1. Reetz's algorithm searches the LPC spectra for a prominence below 400 Hz which is constant in frequency and is present for at least 40 msec. Reetz tested his algorithm on a set of British English vowels and vowel fragments for which perceptual judgments on the presence of nasality were available. Overall, the results were quite good: the algorithm classified the vowel tokens in a manner very similar to the judgments of nasality by the listeners.

This metric is relatively straightforward to implement, and it could be used to detect the presence of nasalization. However, assuming that this algorithm identifies a resonance of a nasal sinus, which will not vary in frequency with changes in velopharyngeal opening, this measure will not provide information about differences in degree of nasalization. Hence it could not be used to derive a detailed picture of changes in nasalization over time. Another limitation of this algorithm is that the LPC analysis had to be done token by token, to determine the optimal number of LPC coefficients in each case, according to criteria which are not discussed in Reetz's paper.

Maeda (1982) describes a spectral metric that quantifies the distribution of spectral energy (flatness) in the F1–F2 region, between 200 and 1700 Hz. The metric was evaluated according to how well its classifications of synthetic French vowels matched perceptual judgments of degree of nasalization of these vowels. If aided by correction for vowel-dependent spectral tilt (the rate, for a particular vowel, at which harmonics decrease in amplitude as frequency increases), the measure captured the gross patterns, though not some of the finer details, of the perceptual data. Recently, Maeda (this volume) described another spectral measure that is more successful at characterizing properties of synthetic vowel spectra that relate to perceptual judgments of nasality. The vowel classifications provided by the metric closely match the data on perceived degree of nasalization for the same tokens, including fine token-specific variations. The metric characterizes spectral flatness in a more indirect fashion than did Maeda's earlier one: this one computes the distance (in Bark) between two spectral prominences in the F1–F2 region of an auditory spectrum. Maeda does not present the details of the algorithm for deciding which two prominences are used, but examination of his figures reveals that they are generally the two strongest prominences below about 12 Bark. Often, these are the two prominences lowest in frequency. When there are not two peaks which are clearly the most prominent (as in some tokens of /a/), then it appears that using the two which are lowest in frequency yields results that correlate well with perceptual judgements of nasal quality.[11] Defining the relevant spectral prominences in this way provides a metric which is sensitive to vowel identity, which may explain the striking success of the measure in characterizing fairly small changes in observed and perceived nasalization.

Either of Maeda's measures might be used to identify the presence of nasalization in an acoustic signal. Furthermore, since the latter measure detects fairly small changes in nasalization, it could be used to identify increases and decreases in degree of nasalization, thus providing an indication of the time course of nasalization during a vowel. It should be noted that this particular measure, which is designed to correlate with perceptual judgments, will be most reliable for detecting changes in nasalization at small to moderate amounts of velopharyngeal opening. As shown in Maeda's figures (this volume), beyond a certain amount of velopharyngeal opening, perceptual judgments and the spectral measure tend to saturate, showing little change (note that the point at which this happens varies to some extent by vowel). Thus while this latter metric could provide a useful outline of the overall time course of nasalization on vowels, it would provide less reliable information about the amount of nasalization at moderate-to-large degrees of velopharyngeal opening.

4.4. Combining Aerodynamic and Acoustic Techniques

In addition to the combinations of data already discussed, one combination which includes nasal airflow, intranasal pressure, and intrapharyngeal or intraoral pressure along with an audio recording may be of particular value, especially when physiological investigation is not feasible. Warren (1964a, 1964b) describes one such setup for use in a clinical setting (see also the discussion in Baken, 1987); Ladefoged (1991) describes a related setup for use in linguistic fieldwork. The audio signal from this combination of recordings aids in segmentation and is available for spectral analysis. The pressure and airflow data taken together can also be used to estimate velopharyngeal port area via the hydrokinetic equation (e.g., Warren & Deveraux, 1966; Warren & DuBois, 1964). Such calculations are most accurate when airflow is steady or changing minimally (Smith & Weinberg, 1980), so while this method can provide excellent data on velopharyngeal area at specific timepoints, it could not be used to derive a continuous record of changes in velopharyngeal area during running speech. However, the nasal and oral pressure data can be used to derive a time-continuous representation of nasalization. The interpretation of the results is aided by the nasal airflow recording, which can help determine whether changes in nasalization are due to greater velopharyngeal port opening (signaled by increased nasal airflow) or are due to conduction from the oral cavity (signaled by no increase in nasal airflow).

5. GENEREAL EXPERIMENTAL ISSUES

In the preceding sections, we have discussed the different kinds of data that may be gathered on nasalization and velopharyngeal function and the instruments used

to gather such data. In most cases, the raw data represent movement of an articulator or an air mass which has been transduced into an electrical signal. (In the case of electromyography, the raw data *are* bioelectric signals.) The electrical signal itself must then be subjected to a number of manipulations in order to yield data from which appropriate quantitative or qualitative judgments can be made. In this section we briefly discuss these manipulations in order to provide a more complete picture of the investigator's task. The kinds of manipulations discussed here are not specific to studies of nasalization or velopharyngeal function but rather apply to the processing of speech production data in general. It should be noted at the outset that this is by no means a comprehensive treatment of these operations. For further information, the reader is referred to Abbs and Watkin (1976), Baken (1987, especially Chaps. 2, 3, and 10), and Kay et al. (1985).

5.1. Display

An important part of any experimental study is the ability to obtain some visual representation of the data signal. Commonly used display devices include the oscilloscope, the chart recorder (for a paper record of the signal), and on-screen computer displays (from which hard copies may be obtained). It is possible to view the signal *on-line,* that is, as it is being obtained, or to view it at a later time, perhaps for further analysis. On-line monitoring is important if the signal will be used to provide feedback in a clinical setting. It is also important in any experiment for checking the performance of the recording system before and during the experiment itself. More permanent visual records of data signals are often important in identifying patterns in the data.

5.2. Storage

Although the output of a chart recorder or printouts (or hard copies) from a computer may serve as a useful record of the data, it is usually desirable to store the data in a form that allows for computational manipulation and measurement. While some experimenters prefer to record data signals directly to a computer hard disk, others prefer to record the signals to tape or some other recording medium first, subsequently transferring the signals to computer disk. The choice of a storage device is limited by the nature of the signal or signals to be stored.

Rapidly changing (or AC) signals can be stored on analog or digital tape, on a compact disk, or directly on the hard disk of a computer. Slowly changing (or DC) signals, such as those reflecting articulator movement, can be recorded on tape using an FM tape recorder or, if voltage levels are converted to different frequencies, using an analog tape recorder. (Recording multiple channels of data on tape requires an appropriate multichannel tape recorder.) DC signals can also be stored on digital tape or a computer hard disk. In order to store AC or DC signals on a digital device (digital tape or computer hard disk), the signals must be converted first to digital words via a processs of pulse-code modulation.

When inputting a signal to a computer (whether directly or from another storage device), one must attend to the limits of the various storage spaces used and to the number of input channels. The digitizing software and hardware devices will impose some limit on the amount of data that can be input at any one time (e.g., the number of seconds' worth of data which may be entered). If the data are read into memory before being written to a more permanent storage location, then the amount of memory available in the computer may also be an issue. Similarly, one must ensure that there is sufficient hard disk space available for the data.

5.3. Amplification

Amplification of the electrical signal is a crucial step in signal processing because the output of most transducers is not strong enough to serve as input for most signal display or storage devices. In addition, some devices, such as tape recorders, add noise to the signal; the data signal must therefore be amplified well above the noise level. Amplification often occurs several times in the course of data collection. Sometimes the term *preamplification* is used to indicate the earliest stage(s) of amplification. An example of preamplification was described in the section on EMG studies. In such studies, the output of the electrodes is first amplified as close to the muscle as possible (i.e., using very short leads), in order to minimize electrical interference with the inherently very weak muscle potentials. The output of the preamplifier is further amplified as required for analysis.

5.4. Filtering

In the acquisition and analysis of data signals, another critical step involves the removal of unwanted components of the signal, a process referred to as *filtering*. Filtering can be accomplished using either hardware devices or computer software (digital filters). Software filters provide more flexibility, in that one can specify the nature of the filter (i.e., it is an equation that is applied to the sampled data), whereas hardware filters have preset characteristics.

All signals must be low-pass filtered before digitization. The mathematics of data sampling are such that frequencies equal to or above one half the sampling frequency cannot be accurately handled during digitization: frequencies at or above the sampling rate appear as low-frequency artifacts in the digitized signal. This frequency "foldover" or "reflection" effect, called *aliasing*, results in distortion of the digitized signal, rendering it useless for most kinds of analysis. Low-pass filters used to avoid such distortion are called *anti-aliasing* filters. (See Rabiner and Schafer, 1978, for a more detailed discussion of aliasing.) As noted, this kind of filtering is necessary to avoid artifacts in digitzed signals. It must occur prior to digitization.

In other cases, filtering is motivated by the desire to simplify inspection and measurement of the signal. In such cases, legitimate components of a data signal may be filtered out to allow the researcher to measure other components more easily. For example, when one is using airflow data to infer information about articulator movement, which involves very low frequency variations, it may be desirable to filter out higher frequency oscillations in the signal such as those caused by vocal tract resonance and vocal fold vibration. This type of filtering can be done with a hardware or software filter.

Filtering is also done when one or more components of the data acquisition system add interference to the signal. This interference can be filtered out with either a hardware or a software filter, as long as its frequency characteristics are known. Thus, for example, high-frequency noise is commonly introduced in the path from transducer to computer. Typically, this noise is out of the range of the signal of interest and can be removed by a low-pass filter or appropriate smoothing algorithm. Understanding filtering is important because inappropriate filtering (or failure to filter when needed) may result in distortion of the data signal.

5.5. Additional Shaping and Processing

There are other kinds of preliminary manipulation of the signal that may be important for certain kinds of data processing. If one is interested in the absolute amount of change, and not in the direction of change, over the course of a signal, it may be useful to transform the signal so that all excursions from zero are represented as positive values, a process called *rectification*. One may also want to smooth a rectified signal, to deemphasize local oscillations in amplitude, making the overall shape of the curve more apparent. One common method of both rectifying and smoothing a signal is to apply an rms function to the signal. Applying this function, the amplitude of the signal at each time point is squared, resulting in all positive values, then these squared values are averaged over certain time periods (which varies with the rms device or algorithm chosen), and the square roots of these values are computed.

There are also occasions when it is useful to invert a signal to make it easier to interpret the data. For example, the output of the Nasometer shows increased nasalization with positive excursions (although the velum is moving down) and decreased nasalization with negative excursions (although the velum is moving up). Inverting the Nasometer signal, as was done for Figure 18b (lower), makes it easier to relate the patterns in the signal to the articulatory changes they reflect.

Another manipulation that is often done involves subtracting signals from one another. For example, it is sometimes necessary to correct for head movement by subtracting a reference signal (e.g., from the Velotrace support rod) from a data signal (from the Velotrace external lever). In other cases, one data signal is subtracted from another to yield a third data signal. In studies of lip and jaw movement, for example, researchers may choose to remove (by means of subtraction) the jaw component from a lower lip signal. If both upper and lower lip movements are obtained, they may also subtract the lower lip y-movement from the upper lip y-movement to obtain lip aperture.

One may also want to transform a signal by differentiating. Differentiating an articulator displacement signal provides the time-varying velocity of the articulator, and differentiating once again provides the corresponding acceleration signal. The derived signals are important for understanding the nature of speech dynamics. As noted by Kay et al. (1985), however, the process of digital differentiation tends to act as a high-pass filter, and so it is recommended that the raw movement signal to be differentiated be smoothed and that the output of differentiation (velocity and acceleration) also be smoothed, taking care to specify an appropriate algorithm.

6. CONCLUSION

This article has presented an introduction to laboratory instrumentation and techniques for the purpose of showing how nasalization and velopharyngeal function can be studied with experimental methods. Such study elucidates the relation between the physical properties of speech and more abstract levels of phonological organization. It also provides insight into the nature of both normal and disordered patterns of speech motor organization. The research described in other articles in this book attests to the importance of experimental investigations of this kind.

A variety of methods for experimental research have been discussed in this chapter, from those that measure muscle activity and movement to those that measure aerodynamic and acoustic effects of movement. There is no question but that researchers have many options for acquiring experimental data on nasalization (or velopharyngeal function).

As readers consider the different methods, we would like to emphasize the importance of comparisons among different techniques and types of measures. There are several reasons for this emphasis: First, combinations or comparisons of techniques (e.g., Velotrace, electromagnetic articulometry, fiberscope) that are thought to provide the same kinds of measurements (e.g., velum height) are important in evaluating the techniques' accuracy and efficacy. Second, combinations of techniques (e.g., EMG, Velotrace, and spectral analysis) that sample different levels of speech organization are important for understanding the relation between muscle activity and movement, and between movement and sound. Third, combinations of techniques (e.g. Velotrace and a Selspot system) or, preferably, the use of a single technique that samples multiple articulators (e.g., electromagnetic articulometry, MRI) are important for understanding how articulation works as a coordinative system: in this case, how the velum fits into more global patterns of gestural organization.

ACKNOWLEDGMENTS

We are grateful for the support of NIH Grant DC-00121 to Haskins Laboratories during the preparation of this article. Fredericka Bell-Berti, Katherine S. Harris, Ignatius Mattingly, Richard McGowan, Patrick Nye, Joseph Perkell, and Ed Wiley provided invaluable comments and suggestions on the manuscript; also Dorothy Ross provided assistance with manuscript preparation, for which we are very grateful.

NOTES

[1] Automated articulatory tracking can reduce some of the time these analyses take, such as using a graphics tablet and pen connected to a computer for graphical input. In such a system, the projector focuses successive frames on the rear of a screen that is also the base of the tablet. The investigator then traces the articulator positions of interest on the tablet. When the tip of a graphics pen comes in contact with the image on the tablet, the coordinates of that position are automatically stored in the computer (see, for example, Alfonso & Baer, 1982). Even with such automation, the frame-by-frame measures remain quite tedious.

[2] In CT scanning, the scanner rotates around the subject as it produces multiple images from different angles. Computer analysis of the set of images creates a composite in the form of a tomograph. CT scans are commonly used in medical imaging. (For a CT image of the vocal tract, see Dixon & Maue-Dixon, 1980, and for further discussion of CT imaging, see Stone, 1991.)

[3] An acoustic gel is usually applied to the surface of the skin and to the transducer to increase acoustic coupling between them.

[4] Still, the use of any X-ray procedure requires some limits on exposure time. At the Wisconsin facility, X-ray microbeam runs were limited to 15-min time periods. In addition, subjects were permitted to return to the facility no more frequently than once a year.

[5] Only individuals without metal fillings or caps could be subjects because the fillings and caps are opaque in the image, much like the pellets themselves.

[6] In speech research, magnetometer systems were first used to study changes in the diameter of the rib cage and abdomen during respiration. (See Hixon, 1972, for additional information.)

[7] The glue does not appear to keep the coil affixed to the velum for more than a half hour (W. Katz, personal communication, 1992).

[8] The actual percentages calculated have been multiplied by a constant to transform them into numbers which are visible on this type of display.

[9] In a normal speaker, the opening must be greater than about 0.2 cm^2 before there will be acoustic excitation of the air in the nasal cavities (K. N. Stevens, 1992, personal communication).

[10] The intended application of their algorithm is improved detection of the neighboring nasal consonants. Mermelstein (1977) describes an algorithm for detecting nasal consonants by identifying nasalized vowel–consonant and consonant–vowel transitions.

[11] However, the data for /a/ suggest that, in determining which spectral peaks are relevant to the measure, the distance between peaks may be the important factor, rather than prominence (though of course the two are not completely independent)—the two peaks used were never within 3 Bark of each other.

REFERENCES

Abbs, J. H., & Nadler, R. D. (1987). *Users' Manual for the University of Wisconsin X-ray Microbeam.* Madison: University of Wisconsin.

Abbs, J. H., & Watkin, K. L. (1976). Instrumentation for the study of speech physiology. In N. Lass (Ed.), *Contemporary Issues in Experimental Phonetics* (pp. 41–75). New York: Academic Press.

Alfonso, P. J., & Baer, T. (1982). Dynamics of vowel articulation. *Language and Speech, 25,* 151–173.

Baer, T., Gore, J. C., Boyce, S., & Nye, P. W. (1987). Application of the MRI to the analysis of speech production. *Magnetic Resonance Imaging, 5,* 1–7.

Baer, T., Gore, J. C., Gracco, L. C., & Nye, P. W. (1991). Analysis of vocal tract shape and dimensions using Magnetic Resonance Imaging: Vowels. *Journal of the Acoustical Society of America, 90,* 799–828.

Baken, R. J. (1987). *Clinical Measurement of Speech and Voice.* Boston: Little, Brown.

Basmajian, J. V. (1974). *Muscles Alive,* (3rd ed.). Baltimore, MD: Williams and Wilkins.

Basmajian, J. V., & Stecko, G. (1962). A new bipolar electrode for electromyography. *Journal of Applied Physiology, 17,* 849.

Bell-Berti, F. (1976). An electromyographic study of velopharyngeal function in speech. *Journal of Speech and Hearing Research, 19,* 225–240.

Bell-Berti, F. (1980). A spatial-temporal model of velopharyngeal function. In N. J. Lass (Ed.), *Speech and Language: Advances in Basic Research Practice,* (Vol. 4, pp. 291–315). New York: Academic Press.

Bell-Berti, F., Baer, T., Harris, K. S., & Niimi, S. (1979). Coarticulatory effects of vowel quality on velar function. *Phonetica, 36,* 187–193.

Bell-Berti, F., & Hirose, H. (1975). Palatal activity in voicing distinctions: A simultaneous fiberoptic and EMG study. *Journal of Phonetics, 3,* 69–74.

Bell-Berti, F., & Krakow, R. A. (1991). Anticipatory velar lowering: A coproduction account. *Journal of the Acoustical Society of America, 90,* 112–123.

Benguerel, A.-P. (1974). Nasal airflow patterns and velar coarticulation in French. In *Speech Communications Seminar Proceedings* (Vol. 2, pp. 105–112). Stockholm: Almqvist and Wiksell.

Benguerel, A.-P., Hirose, H., Sawashima, M., & Ushijima, T. (1977). Velar coarticulation in French: A fiberscopic study. *Journal of Phonetics, 5,* 149–158.

Buck, M. (1954). Post-operative velopharyngeal movements in cleft palate cases. *Journal of Speech and Hearing Disorders, 19,* 288–294.

Christiansen, R. L., & Moller, K. T. (1971). Instrumentation for recording velar movement. *American Journal of Orthodontics, 59,* 448–455.

Clumeck, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics, 4,* 337–351.

Cohn, A. C. (1990). *Phonetic and Phonological Rules of Nasalization.* Unpublished doctoral dissertation, University of California, Los Angeles. (Distributed as *UCLA Working Papers in Phonetics, 76.*)

Cooper, F. S. (1965). Research techniques and instrumentation. *ASHA Reports, 1,* 153–168.

Curtis, J. F. (1970). The acoustics of nasalized speech. *Cleft Palate Journal, 7,* 380–396.

Dalston, R. M. (1982). Photodetector assessment of velopharyngeal activity. *Cleft Palate Journal, 19,* 1–8.

Dalston, R. M. (1989). Using simultaneous photodetection and nasometry to monitor velopharyngeal behavior during speech. *Journal of Speech and Hearing Research, 32,* 195–202.

Dalston, R. M., Warren, D. W., & Dalston, E. T. (1991). Use of nasometry as a diagnostic tool for identifying patients with velopharyngeal impairment. *Cleft Palate–Craniofacial Journal, 28,* 184–189.

Dickson, D. R. (1962). An acoustic study of nasality. *Journal of Speech and Hearing Research, 5,* 103–111.

Dickson, D. R., & Maue-Dickson, W. (1980). Velopharyngeal structure and function: A model for biomechanical analysis. In N. J. Lass (Ed.), *Speech and Language: Advances in Basic Research and Practice* (Vol. 3, pp. 168–222). New York: Academic Press.

Draper, M. H., Ladefoged, P., & Whitteridge, D. (1959). Respiratory muscles in speech. *Journal of Speech and Hearing Research, 2,* 16–27.

Fletcher, S. G. (1970). Theory and instrumentation for quantitative measurement of nasality. *Cleft Palate Journal, 7,* 601–609.

Folkins, J. W., & Kuehn, D. P. (1982). Speech production. In N. J. Lass, L. V. McReynolds, D. E. Northern, & D. E. Yoder (Eds.), *Speech, Language, and Hearing: Normal Processes* (pp. 109–155). Philadelphia, PA: Saunders.

Fritzell, B. (1969). The velopharyngeal muscles in speech: An electromyographic and cineradiographic study. *Acta Otolaryngologica, Supplement, 250.*

Fritzell, B. (1979). Electromyography in the study of velopharyngeal function: A review. *Folia Phoniatrica, 31,* 93–102.

Fujimura, O. (1990). Methods and goals of speech production research. *Language and Speech, 33,* 195–258.

Fujimura, O., & Lindqvist, J. (1971). Sweep-tone measurements of vocal-tract characteristics. *Journal of the Acoustical Society of America, 49,* 541–558.

Fujimura, O., Miller, J. E., & Kiritani, S. (1975, November). *An X-ray Observation of Movements of the Velum and Tongue.* Paper presented at the 90th Meeting of the Acoustical Society of America, San Francisco, CA.

Fujimura, O., Miller, J. E., & Kiritani, S. (1977). *A Computer-controlled X-ray Microbeam Study of Articulatory Characteristics of Nasal Consonants in English and Japanese.* Paper presented at the 9th International Conference on Acoustics, Madrid, Spain.

Glass, J. R. (1984). *Nasal Consonants and Nasalized Vowels: An Acoustic Study and Recognition Experiment.* Unpublished master's thesis, Massachusetts Institute of Technology, Cambridge.

Glass, J. R., & Zue, V. W. (1985). Analysis and recognition of nasal consonants in American English. In *Proceedings of the IASTED International Symposium on Robotics and Automation,* Lugano, Switzerland, June 24–26, 1985.

Harris, K. S. (1970). Physiological measures of speech movements: EMG and fiberoptic studies. *ASHA Reports, 5,* 271–282.

Harris, K. S. (1981). Electromyography as a technique for laryngeal investigation. *ASHA Reports, 11,* 70–83.

Harris, K. S., Lysaught, G., & Shvey, M. (1965). Some aspects of the production of oral and nasal labial stops. *Language and Speech, 8,* 135–147.

Hattori, S., Yamamoto, K., & Fujimura, O. (1958). Nasalization of vowels in relation to nasals. *Journal of the Acoustical Society of America, 30,* 267–274.

Hawkins, S., and Stevens, K. N. (1985). Acoustic and perceptual correlates on the non-nasal–nasal distinction for vowels. *Journal of the Acoustical Society of America, 77,* 1560–1575.

Henderson, J. B. (1984). *Velopharyngeal Function in Oral and Nasal Vowels: A Cross-Language Study.* Unpublished doctoral dissertation, University of Connecticut, Storrs, CT.

Hirano, M., & Ohala, J. (1969). Use of hooked-wire electrodes for electomyography of the intrinsic muscles. *Journal of Speech and Hearing Research, 12,* 362–373.

Hirose, H. (1971). Electromyography of articulatory muscles: Current instrumentation and technique. *Haskins Laboratories Status Report on Speech Research, SR-25/26,* 73–86.

Hixon, E. (1949). *An X-Ray Study Comparing Oral and Pharyngeal Structures of Individuals with Nasal Voices and Individuals with Superior Voices.* Unpublished master's thesis, University of Iowa, Iowa City.

Hixon, T. J. (1970). An electromagnetic method for transducing jaw movements during speech. *Journal of the Acoustical Society of America, 49,* 603–606 (L).

Hixon, T. J. (1972). Some new techniques for measuring the biomechanical events of speech production: One laboratory's experiences. In *Orofacial Function: Clinical Research in Dentistry and Speech Pathology, ASHA Reports, 7,* 68–103.

Horiguchi, S., & Bell-Berti, F. (1987). The Velotrace: A device for monitoring velar position. *Cleft Palate Journal, 24,* 104–111.

Horii, Y. (1980). An accelerometric approach to nasality measurement: A preliminary report. *Cleft Palate Journal, 17,* 254–261.

Horii, Y. (1983). An accelerometric measure as a physical correlate of perceived hypernasality in speech. *Journal of Speech and Hearing Research, 26,* 476–480.

House, A., & Stevens, K. N. (1956). Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders, 21,* 218–232.

Huffman, M. (1989). *Implementation of Nasal: Timing and Articulatory Landmarks.* Unpublished doctoral dissertation, University of California, Los Angeles. (Distributed as *UCLA Working Papers on Phonetics, 75,* 1990.)

Kahane, J. C. (1982). Anatomy and physiology of the organs of the peripheral speech mechanism. In N. J. Lass, L. V. McReynolds, D. E. Northern, & D. E. Yoder (Eds.), *Speech, Language, and Hearing: Normal Processes* (pp. 109–155). Philadelphia, PA: Saunders.

Katz, W., Machetanz, J., Orth, U., & Schönle, P. (1990). A kinematic analysis of anticipatory coarticulation in the speech of anterior aphasic subjects using electromagnetic articulography. *Brain and Language, 38,* 555–575.

Kay, B. A., Munhall, K. G., Vatikiotis-Bateson, E., & Kelson, J. A. S. (1985). Processing movement data at Haskins: Sampling, filtering, and differentiation. *Haskins Laboratories Status Report of Speech Research, SR81,* 291–303.

Kelsey, C. A., Minifie, F. D., & Hixon, T. J. (1969). Applications of ultrasound in speech research. *Journal of Speech and Hearing Research, 12,* 546–575.

Kent, R. D., Carney, P. J., & Severeid, L. R. (1974). Velar movement and timing: Evaluation of a model for binary control. *Journal of Speech and Hearing Research, 17,* 470–488.

Kent, R. D., & Moll, K. L. (1969). Vocal-tract characteristics of the stop cognates. *Journal of the Acoustical Society of America, 46,* 1549–1555.

Kiritani, S., Itoh, K., & Fujimura, O. (1975). Tongue-pellet tracking by a computer-controlled X-ray microbeam system. *Journal of the Acoustical Society of America, 57,* 1516–1520.

Kollia, H. B. (in preparation). *Functional Organization of Velar Movements Following Jaw Perturbation.* Unpublished doctoral dissertation, City University of New York.

Krakow, R. A. (1989). *The Articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velar Gestures.* Unpublished doctoral dissertation, Yale University, New Haven, CT.

Kuehn, D. P. (1976). A cineradiographic investigation of velar movement in two normals. *Cleft Palate Journal, 13,* 88–103.

Kuenzel, H. J. (1977). A photo-electric investigation of velar height during vowel articulation: First applications of the Velograph. Universität Kiel, Institut fur phonetik, *Arbeitsberichte,* 8.

Ladefoged, P. (1991). Computerized phonetic fieldwork. *UCLA Working Papers in Phonetics, 78,* 1–6.

Larson, P. L., & Hamlet, S. L. (1987). Coarticulation effects on the nasalization of vowels using nasal/voice amplitude ratio instrumentation. *Cleft Palate Journal, 24,* 286–290.

Lindqvist-Gauffin, J., & Sundberg, J. (1976). Acoustic properties of the nasal tract. *Phonetica, 33,* 161–168.

Lubker, J. (1968). An electromyographic-cinefluorographic investigation of velar function during normal speech production. *Cleft Palate Journal, 5,* 1–17.

Lubker, J., & Moll, K. L. (1965). Simultaneous oral-nasal air flow measurements and cinefluorographic observations during speech production. *Cleft Palate Journal, 2,* 257–272.

Maeda, S. (1982). Acoustic correlates of vowel nasalization: A simulation study. *Journal of the Acoustical Society of America, 72,* S102.

Matthies, M. L., Perkell, J. S., & Williams, M. J. (1991). Methods for longitudinal measures of speech nasality in cochlear implant patients. *Journal of the Acoustical Society of America, 89,* 1960(A).

McClean, M. (1973). Forward coarticulation of velar movements at marked junctural boundaries. *Journal of the Acoustical Society of America, 16,* 286–296.

Mermelstein, P. (1977). On detecting nasals in continuous speech. *Journal of the Acoustical Society of America, 61,* 581–587.

Minifie, F. D., Hixon, T. J., Kelsey, C. A., & Woodhouse, R. J. (1970). Lateral pharyngeal wall movement during speech production. *Journal of Speech and Hearing Research, 13,* 584–594.

Moll, K. L. (1960). Cinefluorographic techniques in speech research. *Journal of Speech and Hearing Research, 3,* 227–241.

Moll, K. L., & Daniloff, R. G. (1971). Investigation of the timing of velar movements during speech. *Journal of the Acoustical Society of America, 50,* 678–684.

Moller, K. T., Martin, R. R., & Christiansen, R. L. (1971). A technique for recording velar movement. *Cleft Palate Journal, 8,* 263–276.

Niimi, S., Bell-Berti, F., & Harris, K. S. (1982). Dynamic aspects of velopharyngeal closure. *Folia Phoniatrica, 34,* 246–257.

Ohala, J. J. (1971, April). *Monitoring Soft Palate Movements in Speech.* Paper presented at the 81st Meeting of the Acoustical Society of America, Washington, DC.

Parush, A. (1984). *Movement Organization in Speech Production: Implications from Studies of Coarticulation.* Unpublished doctoral dissertation, McGill University, Montreal, Canada.

Perkell, J. (1977, December). *On Reducing Radiographic Dosages in Pellet-tracking Experiments.* Paper presented at the 94th Meeting of the Acoustical Society of America, Miami, FL.

Perkell, J., Cohen, M. H., Svirsky, M. A., Matthies, M. L., Garabieta, I., & Jackson, M. T. T. (1992). Electromagnetic midsaggital articulometer (EMMA) systems for transducing speech articulatory movements. *Journal of the Acoustical Society of America, 92,* 3078–3096.

Rabiner, L. R., & Schafer, R. W. (1978). *Digital Signal Processing of Speech Signals.* Englewood Cliffs, NJ: Prentice Hall.

Redenbaugh, M. A., & Reich, A. R. (1985). Correspondence between an accelerometric nasal/voice amplitude ratio and listeners' direct magnitude estimation of hypernasality. *Journal of Speech and Hearing Research, 28,* 273–281.

Reetz, H. (1991). Extracting nasality from speech signals. *Proceedings of the XIIth International Congress of Phonetic Sciences,* Aix-en-Provence, France (pp. 406–409).

Rossing, T. D. (1990). *The Science of Sound.* Reading, MA: Addison-Wesley.

Ryan, W. J., & Hawkins, C. F. (1976). Ultrasonic measurement of lateral pharyngeal wall movement at the velopharyngeal port. *Cleft Palate Journal, 13,* 156–164.

Sawashima, M. (1977). Fiberoptic observation of the larynx and other speech organs. In M. Sawashima & F. S. Cooper (Eds.), *Dynamic Aspects of Speech Production* (pp. 31–46). Tokyo, Japan: University of Tokyo Press.

Sawashima, M., & Ushijima, T. (1971). Use of the fiberscope in speech research. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, University of Tokyo, 5,* 25–34.

Schönle, P. W., Gräbe, K., Wenig, P., Höhne, J., Schrader, J., & Conrad, B. (1987). Electromagnetic articulography: Use of alternating magnetic fields for tracking movements of multiple points inside and outside the vocal tract. *Brain and Language, 31,* 26–35.

Seaver, E. J., III, & Dalston, R. M. (1990). "Using simultaneous nasometry and standard audio recordings to detect the acoustic onsets and offsets of speech," *Journal of Speech and Hearing Research, 13,* 358–362.

Shawker, T. H., Sonies, B. C., & Stone, M. (1984). Soft tissue anatomy of the tongue and floor of the mouth: An ultrasound demonstration. *Brain and Language, 21,* 335–350.

Smith, B. E., & Weinberg, B. (1980). Prediction of modeled velopharyngeal orifice areas during steady flow conditions and during aerodynamic simulation of voiceless stop consonants. *Cleft Palate Journal, 19,* 172–180.

Stevens, K. N., Fant, G., & Hawkins, S. (1987). Some acoustical and perceptual correlates of nasal vowels. In R. Channon & L. Shockey (Eds.), *Festschrift for Ilse Lehiste* (pp. 241–245). Dordrecht, Netherlands: Foris.

Stevens, K. N., Kalikow, D. N., & Willemain, T. R. (1975). A miniature accelerometer for detecting glottal waveforms and nasalization. *Journal of Speech and Hearing Research, 18,* 594–599.

Stevens, K. N., Nickerson, R. S., Boothroyd, A., & Rollins, A. M. (1976). Assessment of nasalization in the speech of deaf children. *Journal of Speech and Hearing Disorders, 19,* 393–416.

Stone, M. (1991). Imaging the tongue and vocal tract. *British Journal of Disorders of Communication, 26,* 11–23.

Stone, M., Shawker, T., Talbot, T., & Rich, A. (1988). Cross-sectional tongue shape during vowels. *Journal of the Acoustical Society of America, 83,* 1586–1596.

Subtelny, J. D. (1956). A laminograph study of nasalized vowels produced by cleft palate speakers. Unpublished doctoral dissertation, Northwestern University, Evanston, IL.

Ushijima, T., & Sawashima, M. (1972). Fiberscopic examination of velar movements during speech. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, 6,* 25–38.

Vaissière, J. (1988). Prediction of velum movement from phonological specifications. *Phonetica, 45,* 122–139.

Warren, D. W. (1964a). Velopharyngeal orifice size and upper pharyngeal pressure-flow patterns in normal speech. *Plastic and Reconstructive Surgery, 33,* 148–161.

Warren, D. W. (1964b). Velopharyngeal orifice size and upper pharyngeal pressure-flow patterns in cleft palate speech: A preliminary study. *Plastic and Reconstructive Surgery, 34,* 15–26.

Warren, D. W., & Devereaux, J. L. (1966). An analog study of cleft palate speech. *Cleft Palate Journal, 3,* 103–114.

Warren, D. W., & DuBois, A. B. (1964). A pressure-flow technique for measuring velopharyngeal orifice area during continuous speech. *Cleft Palate Journal, 1,* 52–71.

Whalen, D. H. (1990). Intrinsic velar height in supine vowels. *Journal of the Acoustical Society of America, 88* (S1), S54.

Zimmerman, G., Dalston, R. M., Brown, C., Folkins, J. W., Linville, R. N., & Seaver, E. J. (1987). Comparison of cineradiographic and photodetector techniques for assessing velopharyngeal function during speech. *Journal of Speech and Hearing Research, 30,* 564–569.