CLASSIFICATION OF NORMAL AND DYSPHAGIC SWALLOWS BY ACOUSTICAL MEANS

BY

LISA J. LAZARECK

A Thesis
Submitted to the Faculty of Graduate Studies
In Partial Fulfillment of the Requirements for the
Degree of

MASTER OF SCIENCE

Department of Electrical and Computer Engineering
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES
*****
COPYRIGHT PERMISSION

**CLASSIFICATION OF NORMAL AND DYSPHAGIC
SWALLOWS BY ACOUSTICAL MEANS**

BY

LISA J. LAZARECK

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of

Manitoba in partial fulfillment of the requirement of the degree

of

MASTER OF SCIENCE

IN ELECTRICAL ENGINEERING

(c) Lisa J. Lazareck, September 2003

ABSTRACT

This paper proposes a non-invasive, acoustic-based method to differentiate between individuals with and without dysphagia or swallowing dysfunction. Swallowing sound signals, both normal and abnormal (i.e., at risk of some degrees of dysphagia) were recorded with accelerometers over the trachea. Segmentation based on waveform dimension trajectory (WDT, a distance-based technique) was developed to segment the non-stationary swallowing sound signals. Two characteristic sections emerged, Opening and Transmission, and 24 characteristic features were extracted and subsequently reduced via discriminant analyses. A discriminant algorithm was also employed for classification, with the system trained and tested using the leave-one-out approach. Overall, 350 signals were used from three bolus consistencies (semisolid, thick and thin liquids). A final screening algorithm correctly classified 13 of 15 control subjects and 11 of 11 subjects with neurological impairments. The proposed method has great potential to reduce the need for videofluoroscopic swallowing studies (current gold standard method for swallowing assessment, which is invasive and non-portable) and the overall clinical assessment of swallowing sound signals.

ACKNOWLEDGMENT

First, I would like to thank my advisor, Dr. Zahra Moussavi, for her steady guidance, commitment, support and enthusiasm. Also, sincere thanks to: Dr. Gina Rempel, for her review, input and assistance in data collection; Dr. Witold Kinsner, for his time and contributions; Dr. James F. Peters, for his constant encouragement and keen interest. In addition, I would like to thank the University of Manitoba and Department of Electrical and Computer Engineering staff for their guidance. Second, this work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada and I thank them for their sponsorship. Finally, I wish to thank my parents, family and friends for their steadfast support, patient understanding and love.

To My Parents:

Striving for Excellence,

I Learned From You.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1   Motivation

Swallowing dysfunction or dysphagia is most common in individuals with neurological impairment such as brain-stem stroke, head/neck injuries and spinal cord injury with anterior cervical fusion [Logemann, 1986]. Also affected are individuals who suffer from some degrees of muscular incoordination and/or paralysis as a result of poliomyelitis (viral inflammation of the gray matter of the spinal cord), Guillian-Barré Syndrome (acute neurological disorder) and cerebral palsy (non-progressive motor disorder, resulting from damage to the central nervous system) [Reddy et al., 1990; Webster's Dictionary, 1988]. Up to 80% of individuals severely affected cerebral palsy have some degrees of dysphagia, placing them at the risk of aspiration (the event of food being drawn into the airway below the glottis) [Selley et al., 1990; Mirrett et al., 1994]. Statistically, dysphagia has reportedly occurred in approximately 13-14%, 30-35% and 40-45% of patients in acute care hospitals, rehabilitation centers and nursing home facilities, respectively [Logemann, 1995]. Repeated choking, malnutrition, dehydration, recurrent respiratory disease, cachexia, and death may all be consequences of recurring aspiration and dysphagia [Reddy et al, 1988]. Despite severe and serious repercussions, our understanding of the fundamental dysphagic swallowing mechanism is still far from complete. Moreover, there is not a well-accepted theory explaining the physiological cause or acoustical characteristics of normal swallowing sounds. Thus, without a

thorough study of the normal swallowing sound, the potential for using acoustical analysis to understand dysphagic swallowing sounds is limited.

The current gold-standard method for the assessment of aspiration is the videofluoroscopic swallowing study (VFSS). VFSS is a radiologic procedure, whereby subjects ingest small amounts of barium-coated boluses while x-rays penetrate the subject and resultant images are video-recorded. VFSS allows immediate visual inspection, detection and localization of abnormalities and demonstrates function impairments, which may result in aspiration [Palmer et al., 1993]. However, VFSS is time-consuming and results in some radiation exposure. The test is not portable and cannot be done in the subjects' typical eating environment, which potentially limits subjects' cooperation. Because of the x-ray exposure and lack of portability, the VFSS cannot be used repeatedly when assessing/monitoring intervention strategies for a patient or assessing an evolving neurological condition of a patient. In addition, the VFSS may be challenging in neurologically impaired children or adults who may not cooperate or position easily into the testing area. Thus, new techniques need to be developed to help assess the performance of the swallowing mechanism.

1.2 Thesis Objectives and Scope

This thesis reports a new, non-invasive method that investigates the fundamental characteristics of both normal and abnormal swallowing sound signals. The particular goals of this thesis are to divide the swallowing sound signal into characteristic segments using waveform dimension trajectory (WDT, a distance-based technique); and to

investigate whether there are any differences between normal and abnormal swallowing sounds. Our hypotheses are:

1. There is a direct relationship between characteristic segments found by the WDT technique and actual stages of swallowing mechanism.

2. There is a difference in sound attributes of each characteristic section between normal and abnormal swallowing sound signals.

The objective of this thesis is to extract the characteristic features to classify between the swallowing sounds of normal subjects and individuals who because of significant neurological impairment are at risk of having food processing and swallowing problems (i.e., dysphagia). In particular, this thesis reports on the techniques for swallowing mechanics corroboration, waveform dimension method utilized for segmentation, feature extraction, modeling, and classification of a number of normal and dysphagic swallowing sound signals. This method may have potential as an alternative assessment to the videofluoroscopic swallowing study, and the overall clinical assessment (screening and diagnosis) of swallowing sound signals.

1.3   Organization of the Thesis

In this Chapter, an introduction to the underlying motivation, objectives and scope of this research has been provided.

Chapter 2 provides a brief assessment of swallowing mechanism and pathology of the swallow, and a comprehensive discussion of swallowing sound signal analysis.

Chapter 3 provides the methodology used in this study. First described are the participating subjects and associated experiments, data acquisition and signal processing

details.   Second described is a new proposed method of adaptive segmentation of swallowing sound signals based on waveform dimension trajectory (WDT).   Third described are the specifics of feature extraction and reduction, and overall feature and signal classification (using the discriminant analysis technique).

Chapter 4 provides detailed results of WDT segmentation, statistical analysis of extracted and reduced features, and overall feature and signal classification performance.

Chapter 5 presents a discussion of the experimental results, including the overall accuracy of the classification model.

Finally, Chapter 6 presents the conclusions of this research and proposes recommendations for future work. Of the proposed method, the limitations are discussed and the contributions are highlighted.

CHAPTER 2

BACKGROUND

2.1   Anatomy of the Swallowing Mechanism

Swallowing is a series of coordinated actions within the pharynx, a tube common to the respiratory and digestive systems. The pharynx is divided into three parts: nasal or naso-pharynx, oral or oro-pharynx and laryngeal pharynx; respectively lying behind the nose, mouth and larynx. The term naso-pharynx is misleading, as functionally the cavity is no part of the pharynx; i.e., the naso-pharynx is entirely respiratory (and made non-collapsible by the attachment of its wall to bone) [Last, 1978]. The oro-pharynx and laryngeal pharynx are both respiratory and digestive, providing a travel path for inspired air and food. Normal swallowing involves intricate control and coordination of three swallowing phases, commonly referred to as oral, pharyngeal and esophageal [Selley et al., 1990; Firmin et al., 1997]. The synchronized pharyngeal actions fall into two types of movements: voluntary and involuntary or automatic [Basmajian, 1970].   In the oral phase, the associated movements are voluntary; whereas, in both the pharyngeal and esophageal phases, the associated movements are automatic.

As illustrated in Figure 2.1 [Last, 1978] at the end of this Chapter, the oral phase begins with the insertion of the bolus (food, solid or liquid form) into the mouth. The food bolus is confined and processed within the mouth by the lips, teeth, tongue muscles and buccinators (special facial muscles that line the cheek and help move food in the mouth to between grinding teeth). Two extrinsic tongue muscles (styloglossi and palatoglossi) raise the sides of the tongue, while a third (genioglossi) depress it in the

midline, transforming the tongue into a chute, where the bolus may slide into the pharyngeal clearance (Figure 2.1a).

The pharyngeal phase begins once the bolus enters the oro-pharynx. The soft palate is elevated or tenses and presses itself against the posterior wall of the pharynx, sealing off the naso-pharynx. At the same time, the larynx, suspended from the hyoid bone by membrane and muscle, elevates with the floor of the mouth to make room for the pharyngeal muscle movement. These muscles also draw the pharyngeal wall over the bolus as a glove is drawn on a finger [Basmajian, 1970]. At this point, both the pharynx and larynx move up to the hyoid bone (Figure 2.1b). The hyoid is a U-shaped bone slung from the skull by ligament and muscle and lies just above the Adam's apple or thyroid cartilage. The bone is strategically situated, in the angle where the front of the neck meets the floor of the mouth, making it the most important landmark in the front of the neck and anchor for tongue and laryngeal cartilages [Basmajian, 1970]. Muscles forming the base of the mouth stabilize the hyoid bone and thyroid cartilage and protect the organs behind them (thyroid gland, larynx, trachea and esophagus).

The esophageal phase begins once the larynx, pharynx and hyoid move up together. The epiglottis is inverted by ary-epiglottic muscle contractions (i.e., those muscles connected or associated to the epiglottis) and closes like a lid over the larynx suspending respiration. As a result, the nose and larynx are sealed off and the bolus has nowhere to go but down the pharynx (Figure 2.1c). By the elevation of the pharynx to receive it, the bolus is much hastened in its descent. The pharynx acts like a snake darting upwards, open-mouthed for its prey – the bolus [Last, 1978]. The outer and inner pharyngeal muscles (constrictor and longitudinal muscles) begin the action of peristalsis,

where waves of contraction pass the bolus inch-by-inch along the length of the digestive tube [Basmajian, 1970]. The larynx, pharynx and hyoid move down together followed by further movement of the larynx and pharynx from the hyoid bone. Within the esophagus, bolus is safe from regurgitation and the laryngeal inlet opens to resume respiration. The epiglottis recovers its normal position by its own elastic recoil, as its framework consists of elastic cartilage. In addition, the crico-pharyngeus muscle, which acts as a sphincter at the lower end of the pharynx, relaxes [Firmin et al., 1997]. The sphincter is always closed, except for the momentary relaxation during deglutition [Last, 1978]. The bolus proceeds on its leisurely way towards the stomach propelled by the slow worm-like contractions of smooth esophagus muscle [Last, 1978]. Lastly, the overall swallowing anatomy is protected behind by the bodies of the upper six cervical vertebrae from which it is separated by loose areolar tissue that permits movement of the tube on them [Basmajian, 1970].

2.2   Diagnosis of Dysphagia

Individuals suffering from some degrees of dysphagia lack the consistent and requisite coordination of the swallowing mechanism. Inefficient oral and pharyngeal clearance of swallowed material results in obvious abnormalities in food processing and subsequent delays in initiating swallows [Casas et al., 1995]. Also, a myriad of possibilities exist for the disturbance of the swallowing mechanism, including anatomical malfunction, paralysis or absence. Thus, both healthy physiology (organs, muscles, etc.) and systemic coordination (of swallowing mechanism) are imperative for normal swallows to transpire.

In addition to deficiencies in swallowing mechanism, other causal roots or sources of dysphagia have been suggested in accordance to the study of a wide-range of maladies. For example, Reddy and associates suggested that dysphagia develops from lesions in certain areas of the cortex and brain stem that control the swallowing mechanism, or from damage to the cranial nerves associated with the swallowing dysfunction. Their studies included stroke patients, head injured patients and others with paralyzing neurological diseases [Reddy et al., 1988]. Leopold and Kagel proposed that dysphagia stems from a pre-oral or anticipatory stage of ingestion, which includes interaction of pre-oral, cognitive, psychosocial and somataesthetic elements engendered by the meal at the time of feeding. Their developed five-stage paradigm of deglutition (pre-oral, preparatory, lingual, pharyngeal, esophageal) was discussed as related to patients with cortical and basal ganglia (of the brain) and psychogenic diseases [Leopold & Kagel, 1997 (a)]. In another study, Leopold and Kagel described the abnormalities leading to dysphagia in patients with Parkinson' disease (illness of later life caused by degeneration in the basal ganglia and characterized by a rhythmic tremor and muscular rigidity [Webster's Dictionary, 1988]) [Leopold & Kagel, 1997 (b)]. During pharyngeal ingestion, they reported impaired bolus motility and stasis, and deficient epiglottic positioning (and subsequent glottic aspiration) abnormalities. During esophageal ingestion, they reported bolus stasis, delayed transport, redirection and esophageal contraction abnormalities. Finally, during lower esophageal sphincter function, they reported opening abnormalities and gastroesophageal reflux [Leopold & Kagel, 1997 (b)]. In addition, Leopold and Kagel described their patients as tachyphagic, an aberration in the pre-oral stage where food accumulates in the mouth, mastication is slowed, and swallowing is delayed and

hesitant [Leopold & Kagel, 1997 (a), (b)]. Most recently, Fillion and Kilcast conjectured that dysphagia, in the form of poor oral sensory integrity and consequently bolus formation and movement, stems from age-related changes in muscle strength and a generalized decrease in epithelial sensitivity (caused by a decrease in the number of nerve endings with increased age). Their studies included healthy adults (ages 20-94) [Fillion & Kilcast, 2001]. On the other hand, Mendell and Logemann hypothesized that dysphagia stems from structural and/or sensory changes in the pharynx. They suggested that symptoms in one anatomic area might be the actual result of a disorder in a different part of the system (structural); and tissue sensitivity and/or damage in the pharynx might affect swallowing coordination (sensory). Their studies included GERD or gastroesophageal reflux disease patients [Mendell & Logemann, 2002]. Overall, it is the delicate nature of the swallowing mechanism that lends itself to increased vulnerability and possible dysphagia.

2.3 Swallowing Sound Signal Analysis

Some reproducible characteristic sound patterns have been reported to be heard during auscultation of swallows with a stethoscope [Dempsey et al, 1990]. In fact, cervical auscultation has been used for swallowing sound analysis as early as 1907, with A.F. Hertz's evaluation of the pharyngeal swallow and secretion accumulation in bulbar poliomyelitis [Vice et al., 1990]. With cervical auscultation, clinicians learn to recognize certain sounds; for example, the sounds of laryngeal stridor and changes in vocalization, or bubbling sounds overlaying respiration, may reflect the presence of excessive secretion in the larynx, and any respiratory clearing actions (coughing, clearing, stridor sounds)

may reflect pharyngeal content penetrating the larynx or trachea [Vice et al.; 1990, 1995]. Stridor is a harsh, high-pitched whistling sound, definitively produced in breathing by an obstruction in the bronchi, trachea or larynx [Webster's Dictionary, 1988].

In recent years, cervical auscultation has been used by clinicians and researchers who investigate acoustical patterns and temporal relations, which exceed auditory perception, at both subjective levels via stethoscope and objective levels via surface microphones or accelerometers (with subsequent computer programming and pattern recognition techniques). Optimal site for sound detection, and best-suited types of acoustic detector unit and adhesive (for the attachment of the detector) have been thoroughly investigated for acoustical analysis of the pharyngeal swallow [Takahashi et al., 1994 (a); Cichero & Murdoch, 2002].

The research team of J.F. Bosma of Maryland has given major contributions to the field. In the early 1990's, Bosma and his team identified three distinct and previously unrecognized structures within swallowing sound signals of infants: Initial Discrete Sound (IDS), Bolus Transit Sound (BTS) and Final Discrete Sound (FDS), where the discrete sounds demonstrated the remarkably stable repetition of the swallow composite [Vice et al., 1990]. Hamlet and her co-workers also documented the characteristic swishing double-click (IDS and FDS) as the bolus passes through the pharynx into the esophagus [Hamlet et al., 1994]. Coincidentally, and referenced in another early 1990's work [Selley et al., 1990], the two 'clicks' have been documented as early as 1965 in [Lear et al., 1965]. In a subsequent 1995 paper, Bosma and his team reported the correlation of cervical auscultation sounds with physiologically recorded swallows in normal infants, in addition to a correction regarding the Bolus Transit Sound; i.e. BTS

was actually a latter sound portion of the Initial Discrete Sound or IDS [Vice et al., 1995]. They inferred that these sounds were generated by an abrupt gesture in the musculature involved in pharyngeal swallow and in upper-respiratory actions. However, they also speculated that the detected sounds (acquired by an accelerometer) might have resulted from either perturbation in the lumenal medium, caused by air or fluid, or direct muscle contraction [Vice et al., 1995]. An additional hypothesis, based on cardiac sound propagation theory, surfaced – where swallowing sound generation was analogous to heart sound generation; i.e., via vibrations set up by pump and valve systems [Cichero & Murdoch, 1998; Selley et al., 1998]. For example, it was postulated that the first swallowing sound was generated by a vibration of the vocal tract complex set by a closing laryngeal valve and lingual pump delivering its first movement against the posterior pharyngeal wall [Cichero & Murdoch, 1998]. Overall, further insight on sounds was achieved; however, no certainty was produced as to the mechanisms that generated the specific sound units. Future work involving video imaging (i.e., through radiography) was suggested [Vice et al., 1995; Cichero & Murdoch, 1998].

In the late 1990's, the majority of acoustical swallow studies were mainly concerned with the timing of the swallow within the breath cycle [Tarrant et al., 1997; Moussavi et al., 1998]. Breath and swallow coordination was studied to assess the maturity and competence of swallowing mechanisms [Rempel et al., 1999] along with the rhythmic breath-swallow patterns in preterm infants [Gewolb et al., 2000]. Later, more attentions were paid to the basic characteristics of the swallowing sound and whether it could be used as an indicator of abnormalities [Qureshi et al., 2000; Chau et al., 2002; Lazareck & Moussavi, 2002]. Bosma and his team studied suck and swallows of infants

over the first month of life, observing that rhythmic stability of suck and swallow appeared to establish by the month end. They claimed with knowledge of this normal development, infants at high-risk of swallowing dysfunction could be identified at early stages [Qureshi et al., 2000]. In a more recent study, they propose that rhythmic suckle feeding can be quantitatively described (i.e., a maturational pattern of suck-swallow rhythmic integration exists) and compared with patterns from infants with bronchopulmonary dysplasia, BPD [Gewolb et al., 2003]. Lastly, Chau and his associates studied pediatric aspiration sound signals and found that of the 94 waveforms, 62% were stationary over small window sizes and 96% violated normality. Implications were suggested for the possible automated detection of aspiration signals [Chau et al., 2002].

2.4   Chapter Summary

Although many studies have analyzed the swallow, studied the diagnosis of dysphagia and investigated properties of the swallowing sound signal, none have proposed a method to distinguish between normal and abnormal signal groups based on sound characteristics. This Chapter has provided the necessary background required to understand the remaining portion of this thesis, which explores a non-invasive method to classify swallowing sound signals as normal or abnormal. The following Chapter begins this exploration with the description of research methodology.
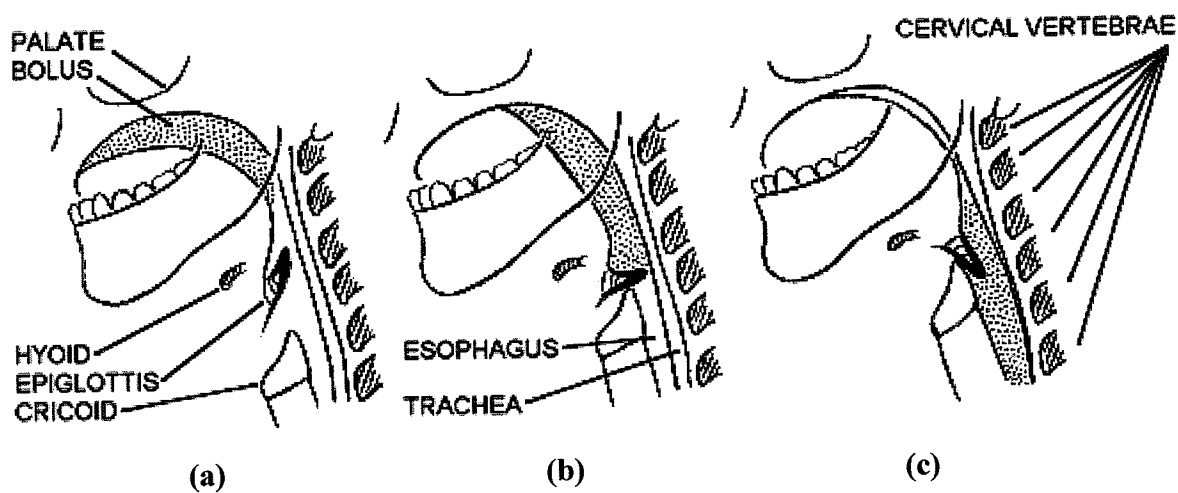
Figure 2.1   Three phases of swallowing: (a) Oral, where the bolus enters the oro-pharynx; (b) Pharyngeal, where the bolus lingers in the oro-pharynx. The larynx has risen, the epiglottis is tilting and yet the hyoid has not moved; (c) Esophageal, where the bolus enters the esophagus. The hyoid and larynx have risen together and the epiglottis is inverted. [Last, 1978].

CHAPTER 3

METHODOLOGY

3.1 Introduction

In Chapter 2, the rudiments of swallowing, dysphagia and subsequent analyses were given. Basic understanding of the swallowing structure gives confidence to proceed with the engineering portion of this thesis, beginning with research methodology. The following sections cover preliminary descriptions of participating subjects and associated experiments, data acquisition and signal processing details. Following are the new proposed method of adaptive segmentation of swallowing sound signals based on waveform dimension trajectory (WDT) and the succeeding approach for feature extraction, reduction and overall feature and signal classification using the discriminant analysis technique.

3.1.1 Terminology

In this study, the term 'normal' was applied to the swallowing sounds recorded from healthy subjects with no history of swallowing dysfunction. These individuals were considered as 'control' subjects. On the other hand, the term 'abnormal' was applied to the swallowing sounds recorded from patients with swallowing disorders. For these subjects, we analyzed the swallowing sounds in which aspiration had not transpired. This selection was chosen for two reasons: first, the recording was stopped as soon as an aspiration occurred. Therefore, there were very few swallows with aspiration (one per patient) for analysis. Second, and more importantly, our goal was to develop a technique

that could identify patients at risk of aspiration, only by acoustical means; otherwise, detection of aspiration is fairly simple as the breath sound immediately following an aspiration changes significantly (i.e., is damped due to an external object in the airway). Therefore, the swallowing sounds analyzed in this study were normal and marginally normal (i.e., 'abnormal,' from our definition above). On the following page, three swallowing sound signals of a normal subject are depicted in both time and frequency (i.e., spectrogram) domains (Figure 3.1). The swallows are interspersed between breath sound signals, both inspiration and expiration, and are respectively situated between 0.6-1.1 msec, 1.7-2.3 msec, and 3.4-3.9 msec. Additionally, and following further explanation in Section 3.4, one typical normal and abnormal swallowing sound signal is illustrated in the time domain (Figure 3.4).
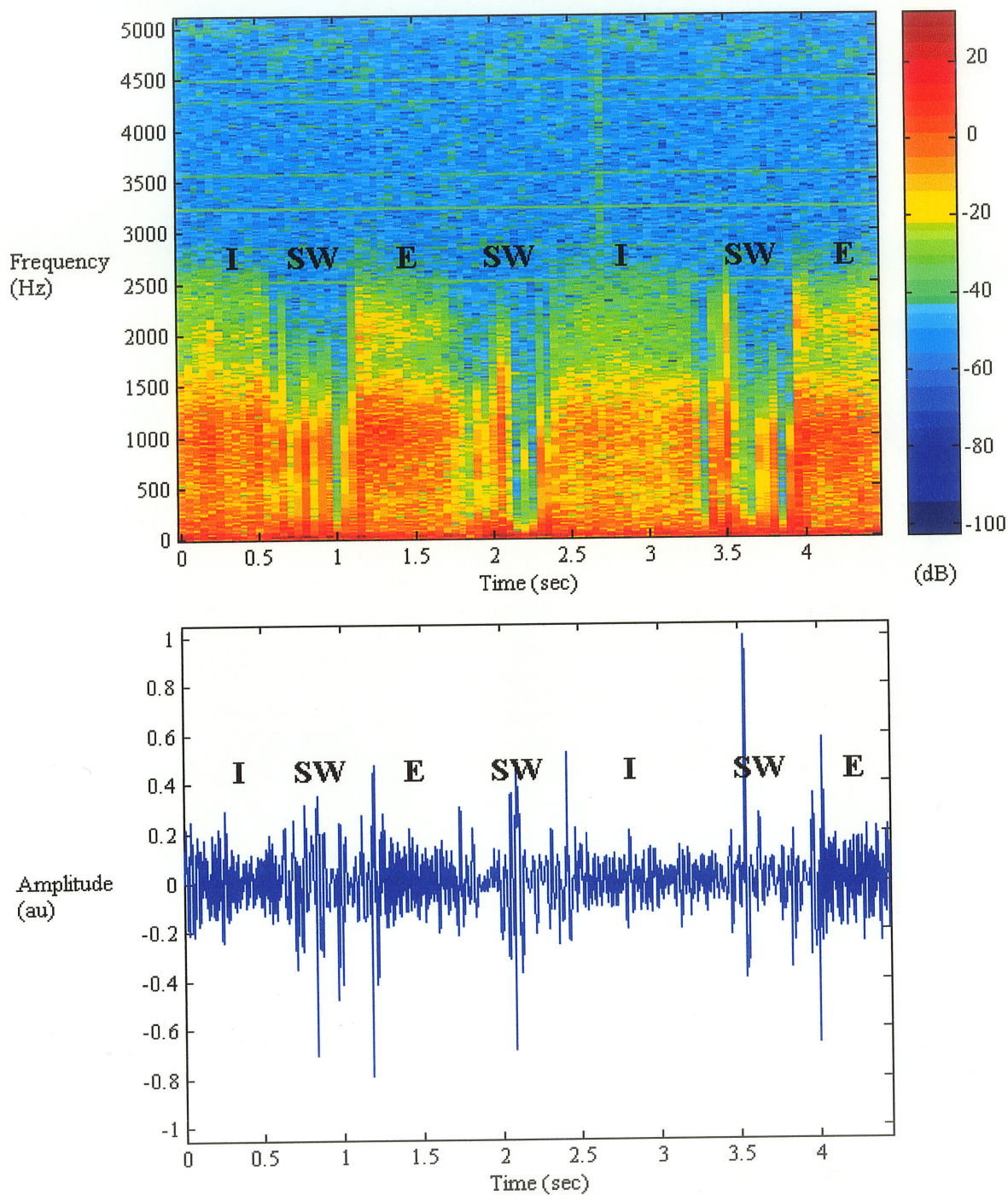
Figure 3.1 Swallowing Sound Signals, Time and Frequency Domains (bottom and top illustrations): three swallows of a normal subject, interspersed between breath sound signals (inspiration and expiration). Legend: 'au,' arbitrary units for normalized amplitude; 'SW'=swallow, 'I'=inspiration, 'E'=expiration.

3.1.2 Subjects

Two groups of subjects participated in the study, including 15 controls and 11 patients. The controls were 12 children (ages 3-16) and three adults (ages 35, 38, 54) with no history of swallowing dysfunction, eating or nutrition problems or lower respiratory tract infections. Controls were tested in the Respiratory Acoustics Laboratory at the Children's Hospital, Winnipeg, Manitoba. The patients were 11 young adults (ages 16-25), divided into two groups: subjects with swallowing dysfunction due to cancer and/or brain tumor (2 subjects, at age 16) and severely cerebral palsy but feeding orally with no clinical signs or symptoms of aspiration (9 subjects). The former group was recruited from the Winnipeg Children's Hospital; they were considered healthy prior to illness (i.e., illnesses were non-congenital). These two subjects were tested in the videoflouroscopy room, Radiology Department (Winnipeg Children's Hospital), during their regular swallowing assessment. The latter group consisted of non-ambulant, wheelchair patients affected by spastic quadriparetic cerebral palsy (Gross Motor Function Scale 4 or 5) [Palisano et al., 1997]. These patients resided in a congregate care setting (St. Amant Centre in Winnipeg, Manitoba). Each of the subjects had good head control, but no independent sitting ability, moderate or severe cognitive impairment, and good nutritional status (as judged by their Centre's clinical dietician). In addition, they were free of lower respiratory tract infections for at least one year prior to investigation. Subject selection was based on the recommendation of their care providers. These subjects were tested in St. Amant Centre, their familiar care environment, during regular meal or snack times, by their regular feeders (the mealtime routine was disturbed as little as possible). In each experiment, informed consent was obtained from all participants or

their parents/guardians prior to the study. The study was approved by the Ethics Committee for the Use of Human Subjects in Research of the University of Manitoba. Lastly, Dr. Zahra Moussavi and Dr. Gina Rempel, University of Manitoba, conducted the studies in 1998. Lastly, it is important to note that the convenience sample of subjects utilized (i.e., controls and heterogeneous mix of subjects with neurological impairments) was not collected for the purpose of this retrospective research.

3.1.3 Experiments

Subjects participated in one of the two conducted experiments: videofluoroscopic swallowing study (VFSS) or audio-recording session (non-VFSS). The VFSS was administered for 3 healthy adults subjects and 2 subjects with cancer and/or brain tumor that were thought to be aspirating based on the clinical evaluation, with both video and audio signals recorded. Audio recording alone (non-VFSS) was performed on 12 healthy children and 9 subjects with cerebral palsy patients, with only audio signals recorded. For both experiments, subjects were fed three textures; 'semisolid,' 'thick liquid,' and 'thin liquid,' in the same order. Respectively, pre-packaged pudding, pudding diluted with an equal quantity of milk, and fruit juice were consumed. The physician in charge monitored each feeding for all subjects. Subjects were offered five to ten 5 mL spoonfuls of pudding and single-bolus sized sips (180 mL in total) of thick and thin liquids respectively. Prior to feeding and subsequent recording, all subjects were examined by the physician via clinical auscultation to confirm their status (i.e., normal, abnormal). Overall, 350 signals were utilized, with specific details illustrated in Table 3.1 at the end of Section 3.1.

3.1.4 Data Acquisition & Signal Processing

In both experiments, breath and swallowing sound and lung sound signals were recorded by accelerometers (Siemens EMT 25 C) being placed over the supersternal notch of the trachea and midclavicular 2nd intercostals space, respectively and were secured with double-sided adhesive tape rings. The signals were amplified and band-pass filtered (50-2500 Hz) to minimize very low-frequency movement artifacts and high frequency noises. During VFSS experiments, breath and swallowing sounds along with the video images were simultaneously recorded with a Super VHS Sony VCR. Later, the audio/video were digitized at an 11.025 kHz sampling rate and 30 frames/second with frame grabber (ATI Multimedia Software ®). Two sample frames, from the VFSS of a normal subject, are shown in Figure 3.2. During non-VFSS experiments, simultaneously with breath and swallowing sounds, airflow was also recorded by nasal cannulae attached to a Fleisch (Number 3) pneumotachograph and Validyne differential pressure transducer (Northridge, CA). The signals were digitized at a 10.240 kHz sampling rate; however, airflow signals were later decimated to 320 Hz.

Swallowing sounds were extracted from each recording. The start and end of each swallow were determined manually by repeated listening and monitoring of the signal in time and frequency (spectrogram) domains. The start point of a swallow was chosen as the first point of the deglutition apnea after any inspiration or expiration. The end point of the swallow was chosen as the sample point preceding the FDS (Final Discrete Sound) if present, or by the breath immediately following deglutition apnea. Each swallow event was subjectively chosen through auditory analysis of the sound signal (both datasets), visual inspection of the signal in the time and frequency domain

(both datasets), and visual inspection of the video frames (VFSS dataset). A time transcription was prepared of the 'Swallow Commencement' and 'Swallow Termination' for each extracted swallow. In addition, for each swallow, another time transcription was prepared for the 'IDS (Initial Discrete Sound) Commencement,' and 'IDS Termination.' Other acoustically-evident modulations were also noted, as most swallowing sound signals included typical onomatopoeically designated 'clacks,' 'gulps,' 'squishes,' and 'braaps.' The overall averaged timings (across normal subjects) were compared to those previously published (see Section 3.3, Feature Extraction).

(a)   (b)
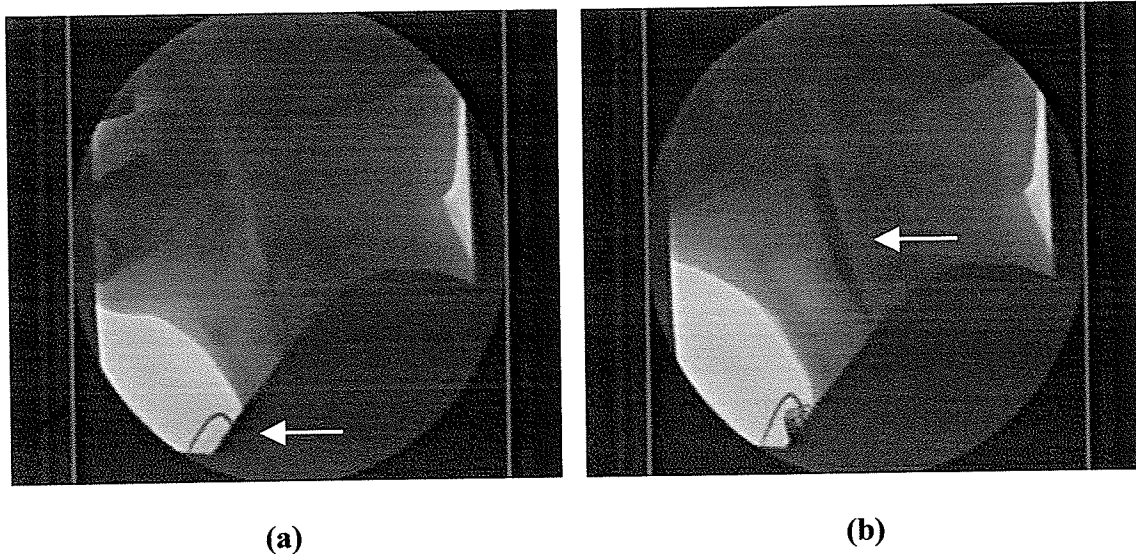
Figure 3.2   Two sample frames from the VFSS of a normal subject: (a) Pre-deglutition, where the mouth opens to accept the bolus (teeth are visible). The white arrow indicates the accelerometer position on the neck; (b) Deglutition, where the darkened bolus is noticeable within the esophagus. The white arrow indicates the bolus as it travels down the throat.

Texture	Total # Swallows	Normal # Swallows	Abnormal # Swallows
Semisolid	108	58	50
Thick Liquid	135	97	38
Thin Liquid	107	96	11
Overall	350	251	99

Table 3.1   Tally of the number of swallowing sound signals acquired per texture, normal and abnormal subjects.

3.2 Adaptive Segmentation Using Waveform Dimension Trajectory (WDT)

3.2.1 Preliminary Investigation of Segmentation Techniques

Given the fact that swallowing sounds are non-stationary by nature, preliminary studies began by analyzing normal swallowing sound signals by dividing the signal into stationary segments using fractal dimension concept [Lazareck & Moussavi, 2002]. In particular, we used variance fractal dimension (VFD), which is a fractal-based measurement describing the underlying signal complexity or major changes in the signal's variance [Kinsner, 1994]. The advantages of VFD include simplicity and aptitude for real-time fractal analysis. VFD is computed by analyzing the spread of the signal amplitude increments in time domain:

$$VFD = E + 1 - H, \tag{3.1}$$

$$H = \lim_{\Delta t \to 0} \frac{1}{2} \frac{\log[\text{variance}(\Delta S)_{\Delta t}]}{\log(\Delta t)};$$

$$\Delta t = |t_2 - t_1|; \quad (\Delta S)_{\Delta t} = S(t_2) - S(t_1), \tag{3.2}$$

where E is the embedding dimension, H is the Hurst exponent, S is the signal in time domain (swallowing sound in this case), Δt is the length of the time window and ΔS is the change in signal for the associated time window (both Δt and ΔS are dependent on scale; for example, Δt most commonly shrinks towards zero in a dyadic fashion). Embedding dimension is the Euclidean description of any object; a point, line, surface, volume and time-space have dimensions 0, 1, 2, 3 and 4 respectively. In this case, the

swallowing sound signal falls somewhere in between a line and a surface. Hence, it must have a dimension between 1 and 2. The Hurst exponent, named after the scientist who studied droughts along the Nile River in Egypt, relates changes in signal magnitude to time and is a key factor in all scale analysis [Peitgen, et al., 1992]. VFD is calculated for a specified window size, which is being slid through the entire signal with some specified sample overlap. The resultant collection of VFD values is denoted as variance fractal dimension trajectory (VFDT). The spread of VFDT is indicative of the multifractal richness apparent in the signal; for example, a unifractal object yields a flat line (dimension of 1). Further discussion and derivation of VFDT algorithm can be found in [Kinsner, 1994]. With significant changes in VFDT, subsequent stationary divisions were created and labelled as 'click' and 'non-click' according to auditory analysis designations. Then, by studying the average power of the short-time Fourier transform of each segment, distinctive differences were observed between the click and non-click [Lazareck & Moussavi, 2002]. Later, waveform dimension (WD) was employed [Katz, 1998] as a tool for segmentation instead of VFD. Loosely based on the principle of fractal dimension, WD is also a measurement of the degree of complexity or meandering between the points of a signal in time domain. The following section highlights the current tool in detail.

3.2.2 Waveform Dimension Trajectory (WDT)

Waveform dimension (WD) is a term coined by M.J. Katz [Katz, 1998]. WD is a measurement calculated for a specified window size, with the window being slid through

the entire signal creating WDT (analogous to the explanation of VFDT in Section 3.2.1). To begin, the general waveform dimension is:

$$WD = \frac{\log(L)}{\log(d)},$$ (3.3)

where L is the total length of the curve and d is the diameter of the curve or planar extent [Katz, 1998]. For waveforms, 'L' is the total length or the sum of the distances between successive points:

$$L = \text{sum}(\text{distance}(i, i+1)),$$

where dist(a,b) means the distance between two point pairs (i.e. a=(x,y) and b=(u,v)). In addition, 'd' is the planar extent or the diameter of the waveform (that can be considered to be the farthest distance between the starting point (1) and any other point (i) of the waveform):

$$d = \max(\text{distance}(1, i+1)).$$

Next, for discretizing space and for normalizing waveform dimensions, Katz discusses a practical convention, where the standard unit is defined as the size of the smallest convolution or "wiggle" of interest in a shape. Alternatively, and for the final result, Katz utilizes another general convention where the standard unit is considered to be the average step or convolution of a shape. For waveforms, 'a' is the average step or the distance between successive points. Incorporating this convention into Equation 3.3 yields:

$$WD = \frac{\log(L/a)}{\log(d/a)},$$

with,

$$a = \text{mean}(\text{distance}(i, i+1)).$$

Setting,

$$n = \frac{L}{a},$$

or the number of steps in the curve, the final formula for waveform dimension is achieved:

$$WD = \frac{\log(n)}{\log(n) + \log(d/L)}, \qquad (3.4)$$

where (within a time window) n is the number of steps in the waveform [Katz, 1998].

In comparison with other waveform-based algorithms, such as the Higuchi and Petrosian methods, the WDT calculation procedure is reportedly fairly insensitive to noise [Esteller et al., 2001]. Equations 3.2 and 3.4 are similar in construction; however, VFD involves the principle of scaling, whereas the Katz algorithm does not. Overall, we found VFD trajectories heavily dependent on the selected window size and the percentage of (sample) overlap; whereas, WD trajectories were less susceptible to these changes. The WDT was also a faster algorithm than VFDT and resulted in more accurate segmentation (i.e. yielded superior segmented boundaries in comparison with actual physiological events). Thus, we selected WDT for the current study.

3.2.3 WDT Application

An adaptive segmentation method based on the aforementioned waveform dimension (WD) was applied to the 350 swallowing sound signals. For each single swallowing sound extracted from the entire breath and swallowing sound signal, a 512-sample window was used to calculate WD, using Equation 3.4. The window was slid across the entire signal with an increment of one sample at a time, and WD was computed accordingly. The 512-window corresponded to 46.4 and 50 milliseconds for the VFSS and non-VFSS datasets (sampling rates of 11.025 and 10.240 kHz). From studying the swallowing sound signal in time and frequency domains, the shortest characteristic signal of a swallowing sound is a 'click' sound, which at most is approximately 33 milliseconds. Thus, a 512-sample window was long enough to capture the click, and short enough to eliminate any redundant details. Once the window completely slid through the signal, the resultant collection of WD values was denoted as waveform dimension trajectory (WDT), which had a one-to-one correspondence with its analogous swallowing sound signal. The corresponding WD trajectory signals for the typical swallowing sound signals mentioned in Section 3.1.1 are illustrated following Section 3.4 (Figure 3.4). Significant changes in the WDT indicated significant changes in the analogous swallowing sound signal.

From each singular WDT signal, significant changes were marked as the boundaries of characteristic segments of the swallowing sound signal. A significant change was defined as any obvious, albeit subjective, variant within the waveform. For example, Figure 3.3 illustrates an abnormal swallowing sound signal and its corresponding WD trajectory. For the first major WDT contour, time domain boundaries

would be marked between 0.13 and 0.20 msec; whereas, the subsections within the contour lying between 0.13-0.15 msec, 0.15-0.19 msec and 0.19-0.20 msec would not be marked as they are minor variations, relatively speaking. The second and third WDT contours would be marked between 0.30-0.35 msec and 0.38-0.42 msec respectively (Figure 3.3). Major shifts in slope were helpful indicators of significant signal variation. The WDT boundaries were compared to the manually transcribed timings (VFSS and non-VFSS datasets) to investigate whether WD technique was able to detect characteristic sections of a swallowing sound. From the rounds of audio and video assessment, generalized descriptions, as to which sounds generated which WDT peaks, were noted (Figures 3.3 and 3.4).
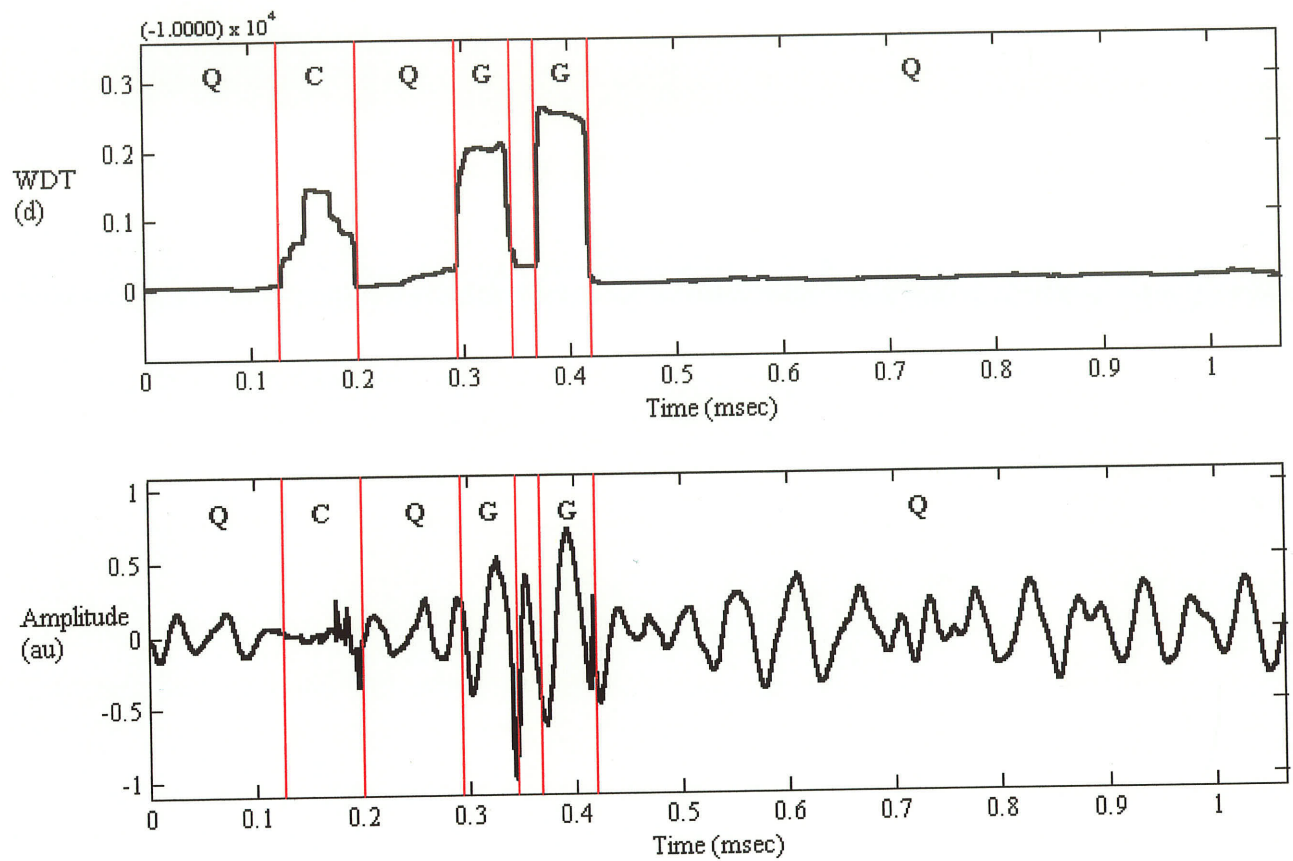
Figure 3.3   Swallowing Sound Signal with Corresponding WD Trajectory: abnormal subject.   Legend: 'au,' arbitrary units for normalized amplitude; 'd,' waveform dimension units [normalized by (WDT−1.0000) x 10^4 for illustrative purposes]; 'C' = click segment, 'G' = gulp segment, 'Q' = quiet segment.

3.3   Feature Extraction

Having done Section 3.2, matching of WDT and time transcription and time transcription with physiological events, the next step was to extract characteristic features for every swallowing section. From the results of Section 3.3, across the 350 signals, two prominent and similar characteristic sections (amongst and between subjects) emerged and were denoted 'Opening' and 'Transmission.' Opening section represented bolus entering the esophagus and was defined as Swallow Commencement to IDS Termination (time transcription titles described in Section 3.1.4). Transmission section represented bolus traveling down the esophagus to the stomach and was defined as the swallowing sound signal from the sample point following IDS Termination to Swallow Termination. Both Opening and Transmission sections were found to be wide-sense stationary, in the sense that the mean value and variance were unchanged per subject. The last part of a swallow, FDS, may or may not be present in a swallowing sound signal. Due to uncertainty of the last section of swallowing sound (FDS), in this part of the study only the Opening and Transmission parts of swallowing sounds were considered to extract characteristic features. Also, note that extraneous noises ('clack,' 'gulp,' and so on) were not removed from the Opening and Transmission sections. The swallowing sound signals of Figure 3.4 have also been demarcated for convenience (see solid red boundary lines), at the end of this Section.

Next, potential features were extracted, all pertaining to Opening and Transmission sections plus the 'Total' swallowing sound signal in its entirety. Stemming from time transcriptions, WD trajectories, and the swallowing sound signal magnitude, three sets of features were computed for each of the Opening, Transmission and Total

sections. The first feature set was 'Time Duration,' or length of swallowing sound signal in the time domain per section. The second feature set was 'Waveform Dimension,' the maximum value of WDT for Opening and Transmission sections and the mean value of WDT for Total section. The third feature set was 'Magnitude,' or mean rectified value of the swallowing sound in time domain for Opening, Transmission and Total sections (normalized by the maximum value of corresponding signal section). Following the initial intuitive features, more were selected and designated for the Opening section only, as Opening was the least contaminated by extraneous noises, such as 'gulps' and 'clacks.' Using Fast Fourier Transform, the power spectrum was calculated for every 100 ms segment of the signal (with 50% overlap between adjacent segments). Then from the spectrum of each 100 ms segment, 'Fpeak,' 'Fmax,' 'Fmean,' 'Fmedian' and 'Pave' were calculated. Fpeak (Hz) was the frequency corresponding to the maximum power spectrum magnitude, Fmax (Hz) was the frequency beyond which the power spectrum magnitude dissipated to less than 10% of its maximum, and Fmean (Hz) and Fmedian (Hz) were the statistical mean and median frequencies of the power spectrum. Pave was the average power calculated over seven specified frequency bands (read subscripts, in Hertz), thus branching into seven additional features: 'Pave$_{150\text{-}300}$,' 'Pave$_{300\text{-}450}$,' 'Pave$_{450\text{-}600}$,' 'Pave$_{600\text{-}750}$,' 'Pave$_{750\text{-}900}$,' 'Pave$_{900\text{-}1050}$,' Pave$_{1050\text{-}1200}$' (dB). Finally, for both Opening and Transmission sections, 'Skewness' and 'Kurtosis' features were computed as follows:

$$skewness = \frac{E[x - \mu]^3}{\sigma^3}, \qquad (3.5)$$

$$kurtosis = \frac{E[x - \mu]^4}{\sigma^4}, \qquad (3.6)$$

where x is the data (in this case, swallowing sound signal in time domain), μ and σ are the mean and standard deviation of x, respectively, and E[ ] represents the expected value of the quantity. Skewness is a third order statistic, measuring the asymmetry of data around its sample mean (Equation 3.5). If skewness is negative, the data are spread out more to the left of the mean than to the right (the opposite is true for positive skewness). Kurtosis is a fourth order statistic, measuring how outlier-prone a distribution is (Equation 3.6). For more outlier-prone distributions, the greater kurtosis will be from three (the opposite is true for less outlier-prone distributions). For Gaussian or normal distributions, skewness and kurtosis are zero and three, respectively, representing data perfectly symmetric in distribution. Some definitions subtract three from the computed value of Equation 3.6 such that the normal distribution has a kurtosis of zero; however, this convention was not adopted for this study [Matlab, 2000].

Following feature extraction, statistical t-tests were performed testing the hypothesis that there was a difference between the two means (normal and abnormal) of each feature. In addition, t-tests were also performed between texture pairs: semisolid-thick liquid, thick liquid-thin liquid, and thin liquid-semisolid. T-test was used to investigate any significant difference, for any of the 24 features, between the normal and abnormal groups ($p < 0.05$).
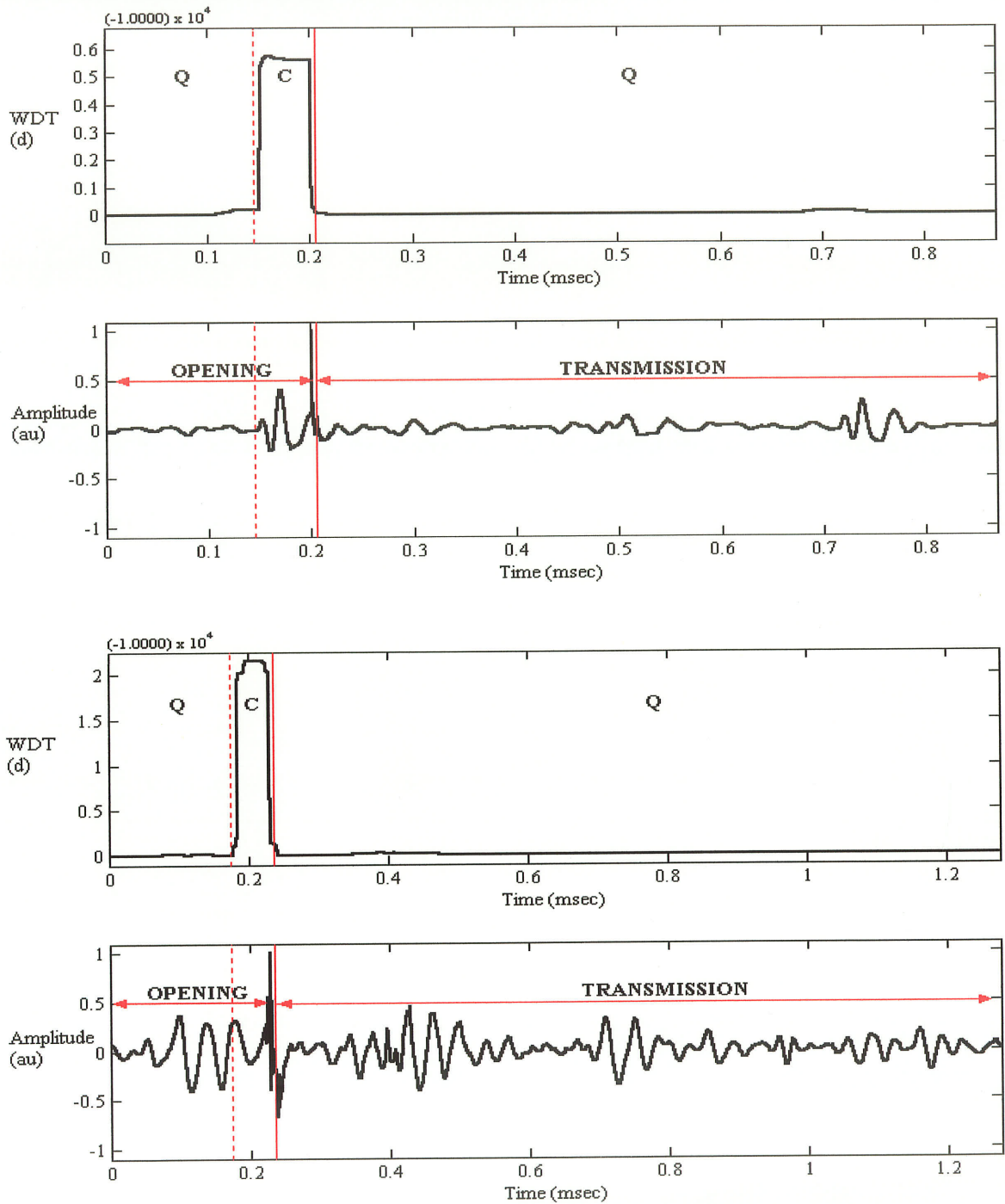
Figure 3.4 Typical Swallowing Sound Signals and Corresponding WDT Signals: normal subject (top), abnormal subject (bottom). Legend: 'au,' arbitrary units for normalized amplitude; 'd,' waveform dimension units [normalized by (WDT−1.0000) x 10^4]; 'C' = click segment, 'Q' = quiet segment; solid red line demarcates Opening and Transmission sections; dotted red line indicates segmentation for matching sound descriptions.

3.4 Feature Reduction & Classification Techniques

3.4.1 Discriminant Analysis

In this study, each swallowing sound signal was classified as normal or abnormal (marginally normal) as described in Section 3.1.1. As each case was known pre-classification, supervised learning was utilized [Bishop, 1995]. The classification model used in this part of study was discriminant analysis, first introduced by Sir Ronald Fisher with emphasis on analyzing variables altogether (as opposed to one at a time) to incorporate important information about their relationships [Norušis, 1994 (a)]. For discriminant analysis, the underlying analysis concept is that linear combinations of selected features are formed and serve as the basis for classifying cases into classes or groups. More specifically, information contained in multiple independent variables may be summarized into a single index. The linear discriminant equation is given as:

$$D = B_0 + B_1X_1 + B_2X_2 + ... + B_pX_p, \tag{3.7}$$

where the X's are the values of the independent variables and the B's are coefficients estimated from the data [Norušis, 1994 (a)]. To distinguish between two groups, the associated D values must differ. The weights (B's) are chosen such that the values of the discriminant function differ as much as possible between the groups. In other words, for the discriminant scores, the ratio (R), defined as:

$$R = \frac{between - groups\ sum\ of\ squares}{within - groups\ sum\ of\ squares}, \tag{3.8}$$

is a maximum. Then, any other linear combination of the variables will have a smaller ratio. The discriminant score for classification was computed based on Bayes' rule. The probability that a case with a discriminant score (D) belongs to a Group (i) is estimated by:

$$P(G_i | D) = \frac{P(D | G_i) P(G_i)}{\sum\limits_{i=1}^{g} P(D | G_i) P(G_i)}, \qquad (3.9)$$

where $P(G_i)$, $P(D|G_i)$, and $P(G_i|D)$ are prior, conditional and posterior probabilities respectively. The group with the largest posterior probability corresponds with the predicted group membership.

3.4.2 SPSS (Statistical Package for the Social Sciences)

Discriminant analysis was applied for feature reduction and classification of the 350 swallowing sound signals using SPSS (Statistical Package for the Social Sciences) Software ®. SPSS is a powerful software package for microcomputer data management and analysis [Norušis, 1994 (a)]. It is so comprehensive and flexible that it can take data from almost any type of file and use them to generate tabulated reports, charts, and plots of distributions and trends, descriptive statistics, and complex statistical analysis [Norušis, 1994 (b)]. The most basic SPSS system includes procedures for frequencies, descriptive and exploratory statistics, cross-tabulation, comparing means (t-tests), simple factorial analysis of variance (ANOVA), linear regression and non-parametric tests [Norušis, 1994 (c)]. Furthermore, additional packages such as the Professional Statistics Package utilized for this thesis, are available that specialize in sophisticated techniques

and procedures (such as discriminant analysis) to measure the similarities and differences in data, classify data, identify underlying dimensions and more. Further details may be found in [Norušis, 1994 (a); Papoulis et al., 2002].

3.4.3 Feature Reduction

To begin with, the 24 computed features were assembled into groups similar in meaning and importance. The following symbols represent the feature groups:

- T = Time Duration & Waveform Dimension,

- M = Magnitude,

- F = Frequency,

- P = Power,

- N = Normality.

In order to select the best group or groups of characteristic features, discriminant analysis was performed with each individual group, utilizing all cases as training data. Thick liquid texture provided 135 swallowing sound signals; this was selected for this part of study as it contained the highest number of overall signals in comparison to thin liquid and semisolid textures. When each group of features was considered as the sole group of features, Group T yielded the highest classification result and therefore was selected as the most important due to this preliminary distinction. Thus, Group T features were included in all subsequent combinations tested. In the second step, all possible combinations of the groups (i.e., T, TM, TMF, ...) were evaluated. The combination with the highest classification accuracy was then selected for further testing. Then, in step three, the same procedure was applied to select the best features within the best

group. For instance, if step two chose 'TMP' as the best combination, step three would produce combinations of the individual features of Group T, Group M and then Group P. For example, Group M contained three features (M_1, M_2, M_3 – representing Opening, Transmission, Total sections) and new combinations for testing would include TM_1P, $T M_1M_2P$, $T M_1M_2M_3P$, and so on. Again, for instance, if TM_1P was selected as the superior selection, the individual features of P would be examined. Finally, the optimal feature set or combination was selected with the plateau of subsequent classification accuracies.

After selecting the reduced feature combination, the dataset was divided into training and testing subsets using the leave-one-out approach. Since the number of subjects was not large enough for randomly dividing between training and testing, the leave-one-out approach was used such that first, the individual swallowing sound signals were grouped together (per subject); second, data of all subjects – except one – was selected as training data to create the classification model. The left out subject's data was selected as testing data to assess the classification model. This procedure was repeated until all subjects were considered as the test subject once. The classification accuracy, however, was calculated considering each swallowing sound as a case. Therefore, for each subject, classification accuracy was calculated for the number of swallowing sounds that the subject contributed. Then, the classification accuracy was averaged between the subjects. The entire procedure of discriminant analysis was repeated for every texture.

3.4.4 Classification Accuracy & Error Rates

This preliminary classification detailed the classification rates of each swallow per subject. From each completed analysis, a subsequent tally of each predicted group membership was recorded. A secondary classification, or screening process, was then employed to assess the overall classification rates for each subject. For each texture and each subject, two error rates were calculated: α, the chance that a swallow was falsely classified as abnormal; β, the chance that a swallow was falsely classified as normal. Each subject had predicted group memberships equal to the number of times discriminant analysis was run; hence, the error rates had to be averaged per subject. Next, the error rates were compared to a threshold of 50%, selected intuitively. If more than 50% of the swallowing signals of a subject were predicted as being normal, the subject's swallowing was considered overall as normal. Conversely, if more than 50% of the swallowing sound signals of a subject were predicted as being abnormal, the subject's swallowing was considered as overall abnormal. After this screening, a new accuracy (or misclassification) was calculated for each texture. For each texture, only subjects contributing more than one swallow were included for the screening procedure; as published reports (regarding symmetry and reproducibility of swallowing sounds) suggest that one should evaluate the acoustic characteristics of swallowing sounds from repeated swallows rather than from one swallow [Takahashi et al., 1994 (b)]. A summary of preliminary and secondary (screening) classifications is provided in the block diagrams of Figures 3.5 and 3.6.

UNKNOWN
SIGNAL

```
┌─────────────────────────────┐      ┌──────────────┐
│ TRAINING:                   │      │ TESTING:     │ ╲
│ All subjects, except Subject # 1. │──▶│ Subject # 1. │  ╲
└─────────────────────────────┘      └──────────────┘   │
                                                         │
┌─────────────────────────────┐      ┌──────────────┐   │
│ TRAINING:                   │      │ TESTING:     │   │
│ All subjects, except Subject # 2. │──▶│ Subject # 2. │   │
└─────────────────────────────┘      └──────────────┘   │
                                                         │
┌─────────────────────────────┐      ┌──────────────┐    ╲    ┌──────────────────┐
│ TRAINING:                   │      │ TESTING:     │     ╲   │ DISCRIMINANT     │
│ All subjects, except Subject # 3. │──▶│ Subject # 3. │     ├──│ CLASSIFIER       │
└─────────────────────────────┘      └──────────────┘     ╱   │ For Each Swallow │
                                                          ╱    │ (Normal, Abnormal)│
┌─────────────────────────────┐      ┌──────────────┐    ╱    └──────────────────┘
│ TRAINING:                   │      │ TESTING:     │   │
│ All subjects, except Subject # 4. │──▶│ Subject # 4. │   │
└─────────────────────────────┘      └──────────────┘   │
              ┊                              ┊           │
              ┊                              ┊           │
              ▼                              ▼          ╱
┌─────────────────────────────┐      ┌──────────────┐ ╱
│ TRAINING:                   │      │ TESTING:     │╱
│ All subjects, except Subject # N. │──▶│ Subject # N. │
└─────────────────────────────┘      └──────────────┘
```

Figure 3.5   Flowchart of discriminant classification, Stage 1.
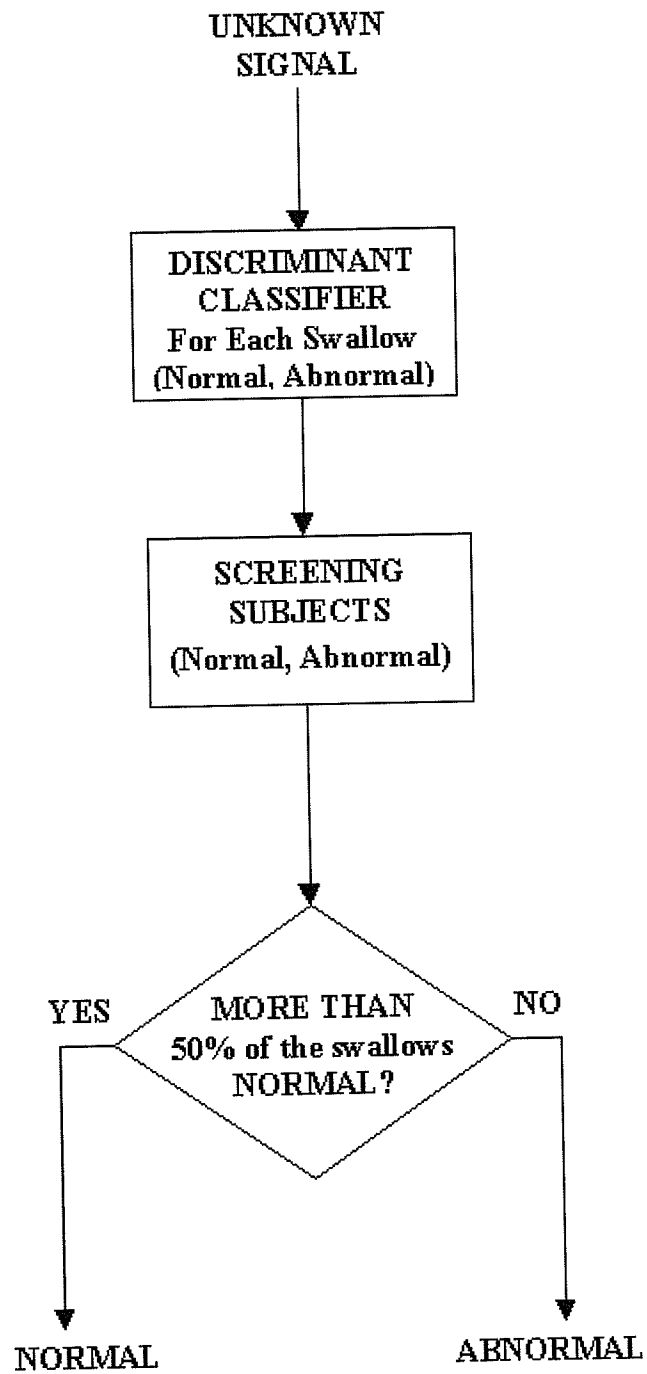
Figure 3.6 Summary of classification procedure, Stages 1 and 2.

3.5   Chapter Summary

This Chapter has discussed the rigorous methodology required to fully assess the swallowing mechanism and swallowing sound signal. First, the preliminary approach was described, including testing of subjects (control and patients), experimental setup (bolus distribution) and overall data acquisition and signal processing (accelerometer and hardware specifications). With the appropriate data recordings, the signals were segmented with ensuing feature extraction and reduction. The proposed segmentation was by means of waveform dimension trajectory (WDT), a simple and quick algorithm. Following were explanations regarding the 24 extracted, grouped and reduced features, overall consisting of: Time Duration, Waveform Dimension, Magnitude, Frequency, Power and Normality. Classification via discriminant analysis was described, with α and β error rates introduced. The application of theses methods to the swallowing sound signals and the results are discussed in the following Chapter.

CHAPTER 4

RESULTS

4.1   Introduction

Adherence to the systemic procedures described in Chapter 3 yielded both intuitive and provocative experimental results. First, swallowing sound correspondence (to swallowing mechanism and WDT segmentation markings) was investigated. All possible venues were exhausted, including time and frequency domains, audio and video data. Second, the underlying characteristics of swallowing sounds were explored through intense feature extraction, and subsequent reduction.   Finally, the outcome of classification for all textures and groups were found. Respectively, the following three sections describe the overall results obtained.

4.2   Adaptive Segmentation Using Waveform Dimension Trajectory (WDT)

Prior to the assessment of WDT segmentation, manual inspection of the swallowing sound signals, by visual and auditory means in time and frequency domains as well as the video data, was crucial, revealing much regarding swallowing sound corroboration with swallowing mechanism. Consequently, these results (described first) were then utilized in the evaluation of WDT results (described second). To begin, the Opening section (defined by Swallow Commencement to IDS Termination) corresponded with the bolus entering the pharynx by way of the opening of crico-pharyngeus muscle. The IDS of the swallowing signal was definitely present in all swallows in controls and subjects (all datasets).   The Transmission section (defined by IDS Termination to

Swallow Termination) corresponded with the bolus traveling down the esophagus and the return of the epiglottis (commonly causing an ensuing 'click' sound, noted as FDS). The FDS of the swallowing signal was noticeably absent in a number of swallows (healthy and dysphagic); therefore, FDS was not included in our analysis. At the end of this Section, Figure 4.1 illustrates a typical swallowing sound and highlights the IDS and FDS. Figure 4.2a depicts the overlay of several swallowing IDS signals to illustrate their morphological similarities; and for illustrative purposes, Figure 4.2b depicts an overlay of several noticeable swallowing FDS signals. The averaged IDS and FDS (where present), for all video-based swallowing sound signals, were recorded at approximately 33.3 milliseconds or one frame (both normal and abnormal subjects). In addition to the IDS and FDS, several extraneous noises were categorized. The 'clack' sound corresponded to tongue movement, and the 'gulp,' 'squish,' and 'braap' sounds were of a more general nature (i.e. not dependent on one physiologically-specific movement). These universal sounds typically emanated from bolus movement within the esophagus (i.e., Transmission section). These superfluous sounds were temporally sporadic and although disruptive (i.e. high in amplitude), were not as prominent as IDS and FDS signals.

Subsequent WDT segmentation revealed a definite correspondence between its subsequent boundaries markings and the aforementioned manual time transcriptions. For example, Figure 3.4 (of Chapter 3: Methodology) depicted the subsequent demarcations of two swallowing sound signals (into Opening and Transmission sections) via WD trajectories in the time domain. First and foremost, the prominent IDS signals (or click segments) were always recognized within the swallowing sound signal as they corresponded with the largest variations or contours in WD trajectory signals. In

addition, extended quiet periods and extraneous noises were also consistently isolated via WDT signals. Figure 3.3 depicts several obvious gulp and quiet segments; whereas, Figure 3.4 depicts quiet segments without any additional noises. A slight recurrent offset was occasionally noticed amongst the 350 demarcated swallowing sound signals (i.e. between WDT and manual time transcriptions), as depicted in Figures 3.3 and 3.4. However, due to the minute resolution of the swallowing sound signal, the offset was deemed insignificant. Overall, manual time transcriptions and WDT segmentation resulted in heavily correlated swallowing sound identification.

(a)
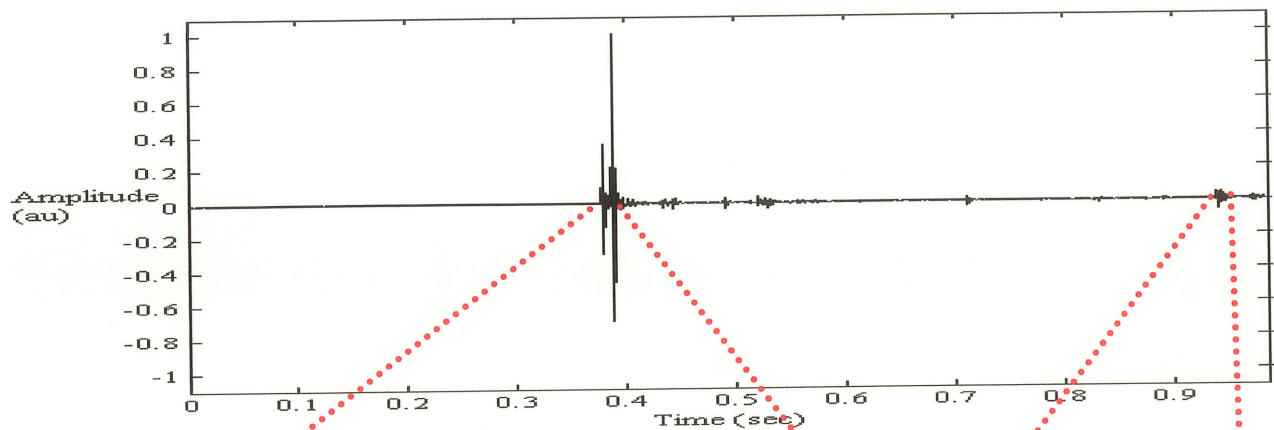
(b)
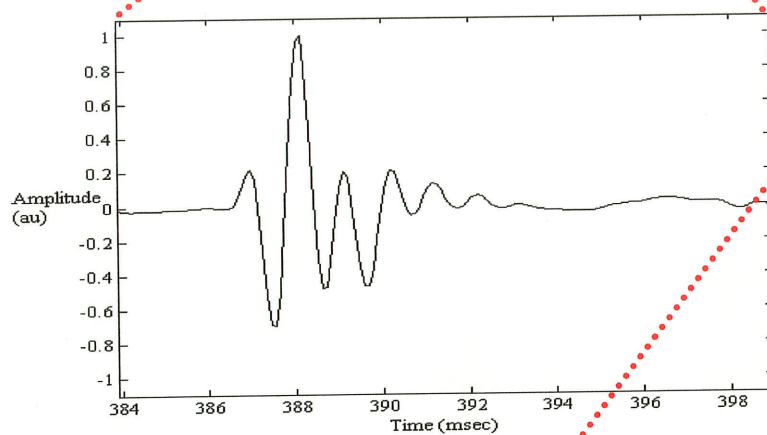
(c)
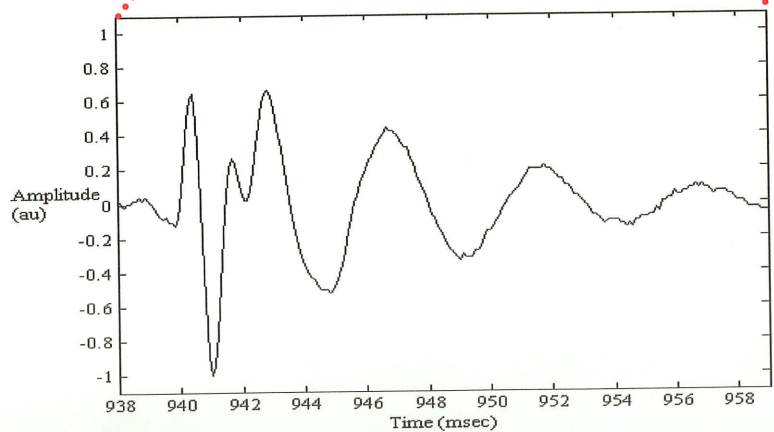
Figure 4.1   Typical Swallowing Sound: (a) signal in its entirety, (b) highlighted and expanded IDS signal, (c) highlighted and expanded FDS signal (although FDS not used in our definition of swallowing sound signal – included here for illustrative purposes). Legend: 'au,' arbitrary units for normalized amplitude.
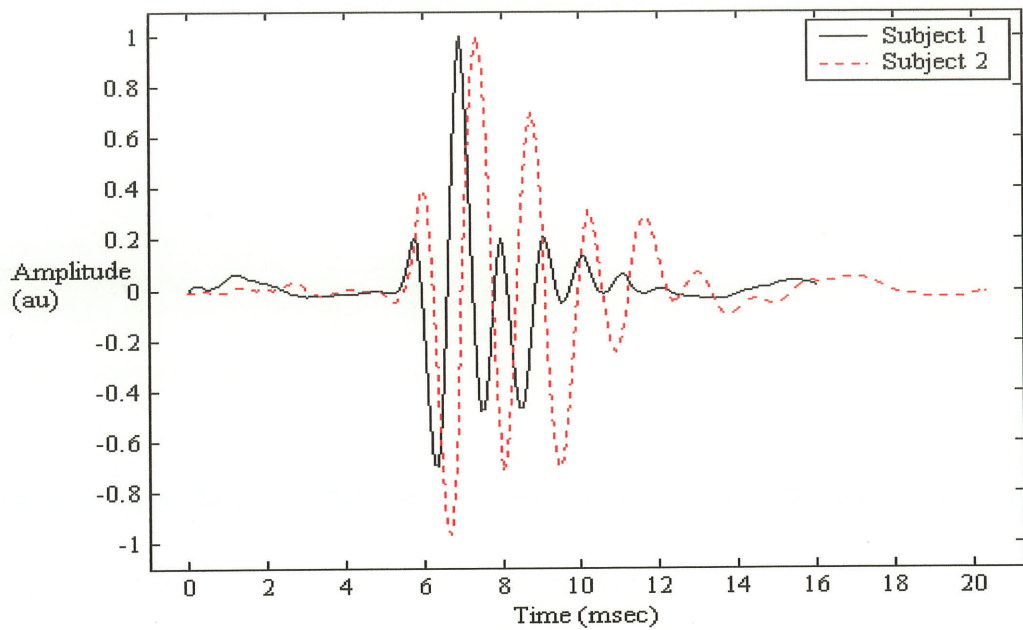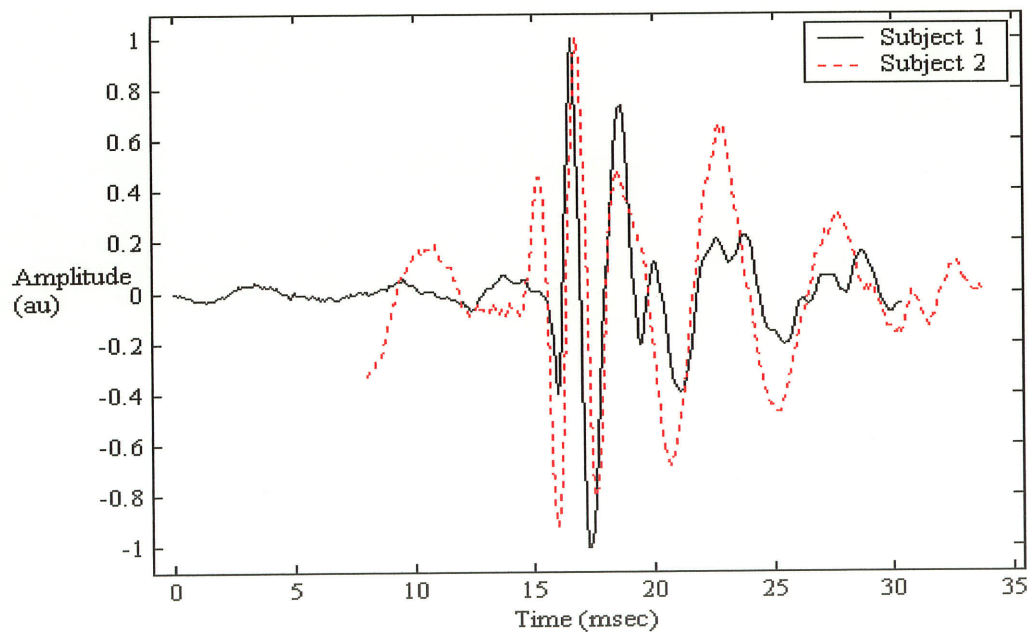
Figure 4.2 Morphological details and overlay of: (a) Initial Discrete Sound or IDS, extracted from two normal swallowing sound signals; (b) Final Discrete Sound or FDS, extracted from two normal swallowing sound signals. Legend: 'au,' arbitrary units for normalized amplitude.

4.3 Feature Extraction

After solid corroboration of WD trajectories and manual time transcriptions, feature extraction and t-tests were performed to identify the most important signal characteristics; i.e., features that would best differentiate between normal and abnormal swallowing sound signals and furthermore, features that would best differentiate between textures for both normal and abnormal swallows. Stemming from manual time transcriptions and corresponding WDT signals, Time Duration was first examined (Table 4.1). The durations of Opening, Transmission and Total sections were significantly longer for abnormal than normal swallowing sounds for both semisolid and thick liquid textures. No significant differences were observed for thin liquid texture. For normal data, Opening Time Duration yielded significant differences between thick and thin liquids textures; whereas for abnormal data, Transmission and Total Time Duration showed significant difference between semisolid and thick liquid as well as between thick and thin liquid textures.

In terms of the Waveform Dimension value, the swallowing sound signals were found to have a dimension between "1" and "2," representing an object between a straight line (one-dimensional object) and a surface (two-dimensional object). Overall, all of the Waveform Dimension values were found to be very close to "1"; thus, the swallowing sound signals were closer to a one-dimensional object in terms of complexity than to a two-dimensional object (Table 4.2). For thick and thin liquids, Opening WDT was significantly higher (i.e. more complex) for normal than abnormal swallowing sounds. Also, thick liquid Total WDT was significantly higher for normal than abnormal swallowing sounds. Between the textures, the differences of Opening WDT between the

thick and thin liquids and between thin liquid and semisolid were significant among normal signals (thick and thin liquids, thin liquid and semisolid), while among abnormal signals it was significant only between thick liquid-semisolid textures. Transmission WDT showed significant differences between semisolid and thick liquid among abnormal swallowing sounds, and finally Total WDT showed significant differences between thin liquid and semisolid among normal swallowing sounds. Table 4.2 shows a summary of the results of WDT for normal and abnormal swallowing sound for different textures.

Magnitude features (of Opening, Transmission and Total sections) yielded significantly higher values for abnormal than normal swallowing sounds, for every texture (Table 4.3). Overall, abnormal swallowing sounds were louder than normal ones. Between the textures, Opening Magnitude showed significant differences between thick and thin liquids among normal swallowing sounds; as well as between thick and thin liquids and thin liquid and semisolid among abnormal swallowing sounds. Transmission Magnitude showed significant differences between semisolid and thick liquid among normal swallowing sounds, and finally Total Magnitude showed significant differences between semisolid and thick liquid among normal swallowing sounds.

Of the Frequency features, for thick and thin liquids, Fpeak was significantly higher for abnormal than normal swallowing sounds (Table 4.4). For semisolid, Fmax was significantly higher for normal than abnormal swallowing sounds. However, Fmean and Fmedian features did not show any significant difference between either normal/abnormal swallowing sounds or between different textures.

Average Power in all seven frequency bands yielded significant differences for semisolid data, where abnormal swallowing sounds had significantly higher average

power than that of normal swallowing sounds (Table 4.5). This result is illustrated in Figure 4.3 (following this text description). In addition, swallowing sound of thin liquid showed significantly higher power ($Pave_{150-300}$) among abnormal than normal swallowing sounds and conversely, higher power ($Pave_{900-1050}$) in normal than abnormal swallowing sounds. Between the textures, all the seven Power features were significant in discernment between semisolid-thick liquid and thin liquid-semisolid textures. Overall, a dampening effect was observed; i.e., the thicker the bolus texture, the lower (or more damped) the average power. This effect can be heard from repeated audition of swallowing sound signals from one subject, changing bolus textures.

Finally, normality features were examined. Considering the results of Skewness, one may conclude that the swallowing sounds are slightly skewed toward either left or right of the sample mean (which is zero as the sounds are recorded by a bandpass filter, without a DC offset). However, since the results of Kurtosis are far above 3, overall one can conclude that the normality is failed in swallowing sound signals. In addition, a subsequent Lilliefors test, for goodness of fit to a normal distribution (for all textures – Opening and Transmission sections), reveals normality is failed in all swallowing sound signals. Nevertheless, while there were no significant differences in the Skewness features between normal and abnormal swallowing sound signals, the Kurtosis showed significant differences between the two groups. On the whole, Kurtosis was larger for normal than abnormal swallowing sound signals. It was also extremely variable between the subjects (i.e., high standard deviations, Table 4.7). Neither Skewness nor Kurtosis showed any significant difference between the textures. Tables 4.6 and 4.7 show a summary of these features between the groups.
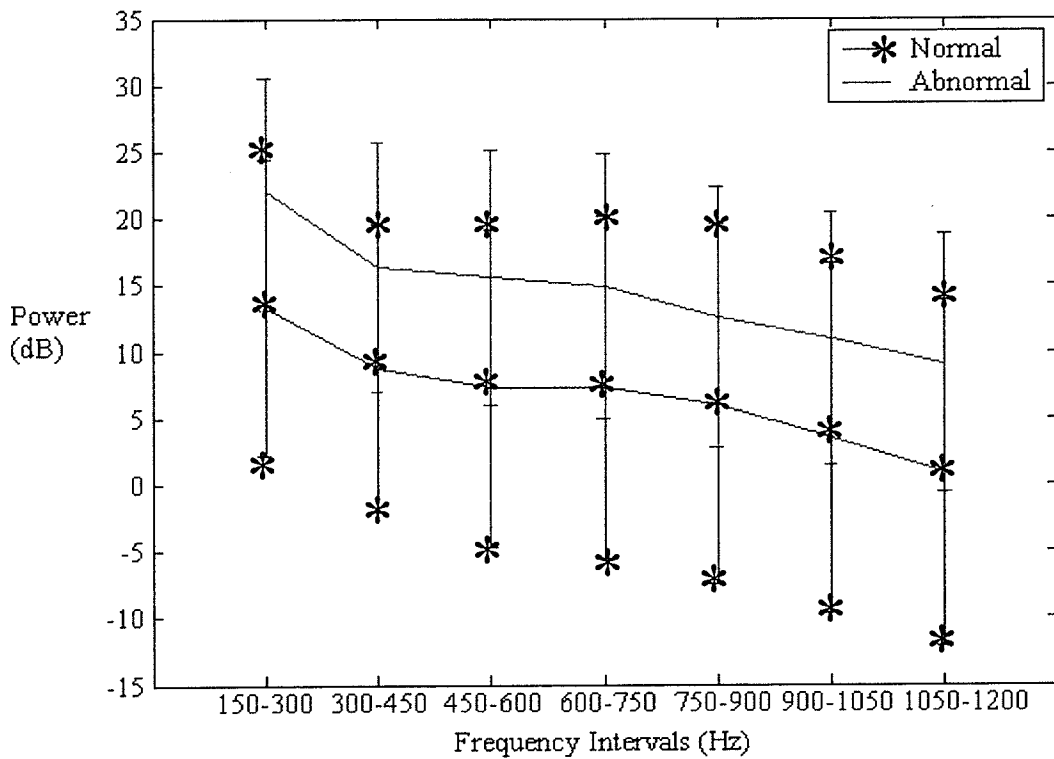
Figure 4.3   Average power in different frequency bands, averaged between the subjects of the two normal and abnormal groups (semisolid data). Error bars represent standard deviation.

Texture	Opening, Duration (s)		SD	Transmission, Duration (s)		SD	Total, Duration (s)		SD
	Normal	Abnormal		Normal	Abnormal		Normal	Abnormal	
Semisolid	0.36 ± 0.30	0.61 ± 0.54	*	0.45 ± 0.17	0.61 ± 0.36	*	0.81 ± 0.34	1.22 ± 0.79	*
Thick Liquid	0.31 ± 0.30	0.59 ± 0.53	*	0.48 ± 0.15	0.79 ± 0.46	*	0.79 ± 0.34	1.38 ± 0.84	*
Thin Liquid	0.43 ± 0.32	0.30 ± 0.30		0.45 ± 0.22	0.54 ± 0.21		0.88 ± 0.39	0.84 ± 0.38	

Table 4.1   Time Duration (seconds): mean ± std of time duration of the Opening, Transmission and Total sections of swallowing sound signals (normal, abnormal) for different textures. Legend: '*' = significant difference (SD) between Normal and Abnormal mean values ($p < 0.05$).

Texture	Opening, WDT (d) (maximum value)		SD	Transmission, WDT (d) (maximum value)		SD	Total, WDT (d) (mean value)		SD
	Normal	Abnormal		Normal	Abnormal		Normal	Abnormal	
Semisolid	1.00024 ± 0.00035	1.00040 ± 0.00057		1.00007 ± 0.00016	1.00005 ± 0.00012		1.00003 ± 0.00003	1.00003 ± 0.00005	
Thick Liquid	1.00034 ± 0.00048	1.00016 ± 0.00022	*	1.00017 ± 0.00034	1.00028 ± 0.00062		1.00003 ± 0.00004	1.00002 ± 0.00005	
Thin Liquid	1.00052 ± 0.00060	1.00013 ± 0.00024	*	1.00009 ± 0.00034	1.00001 ± 0.00002		1.00004 ± 0.00004	1.00001 ± 0.00002	*

Table 4.2   Waveform Dimension: mean ± std of WDT of the Opening, Transmission and Total sections of swallowing sound signals (normal, abnormal) for different textures. Legend: 'd,' waveform dimension units; '*' = same as that of Table 4.1.

Texture	Opening, Magnitude (au) (mean value)		SD	Transmission, Magnitude (au), (mean value)		SD	Total, Magnitude (au) (mean value)		SD
	Normal	Abnormal		Normal	Abnormal		Normal	Abnormal	
Semisolid	0.04 ± 0.04	0.08 ± 0.04	*	0.03 ± 0.03	0.08 ± 0.04	*	0.03 ± 0.03	0.08 ± 0.03	*
Thick Liquid	0.05 ± 0.03	0.08 ± 0.05	*	0.05 ± 0.03	0.10 ± 0.05	*	0.05 ± 0.03	0.08 ± 0.03	*
Thin Liquid	0.03 ± 0.03	0.12 ± 0.06	*	0.03 ± 0.03	0.08 ± 0.03	*	0.03 ± 0.03	0.09 ± 0.03	*

Table 4.3   Magnitude: mean ± std of swallowing sound signal normalized magnitude of the Opening, Transmission and Total sections of swallowing sound signals (normal, abnormal) for different textures. Legend: 'au,' arbitrary units for acceleration, '*' = same as that of Table 4.1.

Texture	Fpeak (Hz)		SD	Fmax (Hz)		SD
	Normal	Abnormal		Normal	Abnormal	
Semisolid	756 ± 253	668 ± 203		2653 ± 620	3141 ± 795	*
Thick Liquid	821 ± 228	707 ± 239	*	2791 ± 596	2719 ± 612	
Thin Liquid	823 ± 298	595 ± 90	*	2724 ± 490	2670 ± 452	

Texture	Fmean (Hz)		SD	Fmedian (Hz)		SD
	Normal	Abnormal		Normal	Abnormal	
Semisolid	1224 ± 191	1330 ± 355		1225 ± 190	1324 ± 352	
Thick Liquid	1278 ± 255	1236 ± 255		1279 ± 251	1234 ± 251	
Thin Liquid	1239 ± 257	1218 ± 167		1241 ± 254	1214 ± 175	

Table 4.4   Frequency (Hz): mean ± std of frequency features of the Opening section of swallowing sound signals (normal, abnormal) for different textures. Legend: * = same as that of Table 4.1.

Texture	$Pave_{150-300}$ (dB)		SD	$Pave_{300-450}$ (dB)		SD	$Pave_{450-600}$ (dB)		SD
	Normal	Abnormal		Normal	Abnormal		Normal	Abnormal	
Semisolid	13.3 ± 11.2	22.1 ± 8.5	*	8.8 ± 10.8	16.4 ± 9.4	*	7.2 ± 11.9	15.6 ± 9.6	*
Thick Liquid	19.6 ± 8.9	20.6 ± 8.4		14.3 ± 9.8	15.8 ± 10.4		13.8 ± 11.0	16.2 ± 12.1	
Thin Liquid	18.0 ± 8.5	24.3 ± 9.9	*	16.2 ± 7.9	18.0 ± 13.0		16.2 ± 8.0	16.5 ± 14.3	

Texture	$Pave_{600-750}$ (dB)		SD	$Pave_{750-900}$ (dB)		SD	$Pave_{900-1050}$ (dB)		SD
	Normal	Abnormal		Normal	Abnormal		Normal	Abnormal	
Semisolid	7.3 ± 12.9	14.9 ± 9.9	*	6.0 ± 13.5	12.5 ± 9.8	*	3.5 ± 13.2	11.0 ± 9.5	*
Thick Liquid	15.9 ± 11.6	14.8 ± 12.0		16.7 ± 12.1	12.7 ± 12.9		15.2 ± 11.9	11.9 ± 13.5	
Thin Liquid	17.6 ± 9.8	14.4 ± 13.4		17.4 ± 10.3	11.4 ± 15.0		15.4 ± 9.7	7.9 ± 15.6	*

Texture	$Pave_{1050-1200}$ (dB)		SD
	Normal	Abnormal	
Semisolid	0.8 ± 13.0	9.1 ± 9.7	*
Thick Liquid	12.8 ± 12.0	10.7 ± 13.7	
Thin Liquid	12.4 ± 9.8	6.7 ± 14.9	

Table 4.5 Power (dB): mean ± std of average power in different frequency bands of the Opening section of swallowing sound signals (normal, abnormal) for different textures. Legend: '*' = same as that of Table 4.1.

Texture	Opening, Skewness		SD	Transmission, Skewness		SD
	Normal	Abnormal		Normal	Abnormal	
Semisolid	-0.1 ± 1.2	0.3 ± 1.1		-0.1± 0.8	0.4 ± 1.0	*
Thick Liquid	-0.2 ± 1.6	0.1 ± 0.5		0.0 ± 1.0	0.2 ± 0.7	
Thin Liquid	-0.2 ± 1.6	0.2 ± 0.7		-0.2 ± 0.7	0.0 ± 0.4	

Table 4.6 Skewness: mean ± std of the Opening and Transmission sections of swallowing sound signals (normal, abnormal) for different textures. Legend: '*' = same as that of Table 4.1.

Texture	Opening, Kurtosis		SD	Transmission, Kurtosis		SD
	Normal	Abnormal		Normal	Abnormal	
Semisolid	27.8 ± 46.6	9.9 ± 10.2	*	14.0 ± 18.9	7.5 ± 7.1	*
Thick Liquid	30.9 ± 62.2	7.4 ± 5.8	*	18.9 ± 28.0	7.7 ± 5.8	*
Thin Liquid	40.9 ± 44.9	5.4 ± 3.0	*	16.9 ± 23.8	4.3 ± 1.8	

Table 4.7 Kurtosis: mean ± std of the Opening and Transmission sections of swallowing sound signals (normal, abnormal) for different textures. Legend: '*' = same as that of Table 4.1.

4.4 Feature Reduction & Classification Techniques

Discriminant analysis was performed with each individual group (all cases as training data), with results tabulated in Table 4.8. As previously mentioned, Group T yielded the highest overall accuracy; therefore, it was included in all further testing. Subsequently, Group T was subdivided into Time Duration and Waveform Dimension features with discriminant analysis performed with each subset. The Time Duration features (first subset) resulted in higher accuracy for semisolid and thick liquid textures in comparison with Waveform Dimension features (second subset). For thin liquid, the accuracies between subsets were comparable. Including Group T in every new group, sixteen new combinations were tested and the results are shown in Table 4.8. Combination 'TMP' and 'TMPN' yielded the highest accuracy: 94.8 % (overall), 94.8 % (within normal class), 94.7 % (within abnormal class). However, since removing the Group N features did not reduce the accuracy of classification, the TMP combination was considered as the optimum feature set. Next, the individual features of Group T, M and P were examined using the leave-one-out approach (as described in Method Section). The eliminated features included Time Duration (of the Total section) as well as $Pave_{150-300}$, $Power_{600-750}$, $Power_{750-900}$ and $Power_{900-1050}$ (of the Opening section). Based on SPSS information, Time Duration Total was removed due to its strong correlation with Time Duration of Opening and that of Transmission features. This result was expected as the sum of Time Duration of Opening and that of Transmission created Time Duration Total. Also according to SPSS, the above mentioned power features were eliminated, as they did not contribute any distinct information or tangible characteristics to the classification model.

Grouping	Overall (%)	Normal (%)	Abnormal (%)
T	84.4	91.8	65.8
M	80.7	83.5	73.7
F	61.5	58.8	68.4
P	72.6	72.7	73.7
N	53.3	39.2	89.5
TM	92.6	92.8	92.1
TF	83.0	86.6	73.7
TP	85.9	86.6	84.2
TN	84.4	91.8	65.8
TMF	93.3	94.8	89.5
TMP	94.8	94.8	94.7
TMN	92.6	92.8	92.1
TFP	90.4	90.7	89.5
TFN	83.0	87.6	71.1
TPN	88.1	86.6	92.1
TMFP	93.3	93.8	92.1
TMFN	92.6	94.8	86.8
TMPN	94.8	94.8	94.7
TFPN	88.9	88.7	89.5
TMFPN	94.1	94.8	92.1

Table 4.8 Preliminary and secondary classification rates for feature reduction. Groups T, M, F, P and N respectively represent Time Duration & Waveform, Magnitude, Frequency, Power and Normality.

With the remaining eleven features, further testing did not reveal any additional classification improvements; i.e., classification accuracy reached a plateau of 95.6 % (overall), 97.4 % (within normal class), 94.8 % (within abnormal class). Overall, the reduced feature set contained the following features: Time Duration (Opening, Transmission), Waveform Dimension (Opening, Transmission, Total), Magnitude (Opening, Transmission, Total), and $Power_{300-450}$, $Power_{450-600}$, $Power_{1050-1200}$. All of these features, with the exception of Waveform Dimension (of the Transmission section), yielded significant differences between normal and abnormal swallowing sound signals for at least one texture.

Discriminant analysis was performed (with the eleven features) for preliminary classification of normal and abnormal swallowing sound signals for each texture. Considering each swallowing sound as one case, overall training and testing classification accuracies were averaged between the subjects (Tables 4.9 through 4.11). For semisolid, thick and thin liquids training accuracies were 82.0%, 94.9%, 94.4% for normals and 89.8 %, 95.1 %, 87.5 % for abnormals, respectively; while the testing accuracies were 85%, 76.6%, 84.6% for normals and 84.8%, 88.1% and 63.3% for abnormals, respectively. Secondary classification or screening followed with overall screening results averaged per subject (Tables 4.9 through 4.11). Applying the screening as the second stage of classification, overall classification accuracy increased significantly for all textures. It yielded correct classifications for 7/9, 10/10, 9/10 of normal subjects and all (10/10, 7/7, 4/4) of abnormal subjects. The only three misclassification cases (two with semisolid, one with thin liquid) stemmed from the two older adult subjects in control (normal) group, ages 38 and 54.

SEMISOLID			
Actual Group	Overall Training (per swallow)	Overall Testing (per swallow)	Overall Screening (per subject)
N	82.0 ± 0.7 %	58.0 ± 10.5 %	77.8 ± 14.7 % (7 of 9 subjects)
A	89.8 ± 0.4 %	84.8 ± 5.6 %	100.0 ± 0.0 % (10 of 10 subjects)

Table 4.9   Training and Testing Accuracy of Classification for Semisolid Texture. Legend: N-Normal, A-Abnormal, mean ± standard error.

THICK LIQUID			
Actual Group	Overall Training (per swallow)	Overall Testing (per swallow)	Overall Screening (per subject)
N	94.9 ± 0.2 %	76.6 ± 8.1 %	100.0 ± 0.0 % (10 of 10 subjects)
A	95.1 ± 0.6 %	88.1 ± 5.2 %	100.0 ± 0.0 % (7 of 7 subjects)

Table 4.10   Training and Testing Accuracy of Classification for Thick Liquid Texture. Legend: N-Normal, A-Abnormal, mean ± standard error.

THIN LIQUID			
Actual Group	Overall Training (per swallow)	Overall Testing (per swallow)	Overall Screening (per subject)
N	94.4 ± 0.4 %	84.6 ± 8.0 %	90.0 ± 10.0 % (9 of 10 subjects)
A	87.5 ± 1.6 %	63.3 ± 8.8 %	100.0 ± 0.0 % (4 of 4 subjects)

Table 4.11   Training and Testing Accuracy of Classification for Thin Liquid Texture. Legend: N-Normal, A-Abnormal, mean ± standard error.

Finally, considering each swallowing sound as a case, two averages were calculated for α and β error rates per subject (Table 4.12). First, 'Exclusion A' described the averaged error rates excluding subjects who contributed one swallow for a texture. Second, 'Exclusion B' described the averaged error rates excluding the two eldest adult control subjects and any subject who contributed only one swallow (one control and three patients). Exclusion A and B averages ranged between 4.0 % to 22.2 % and 0.5 % to 17.2 %, respectively. With the exclusion of adult subjects, Exclusion B averages indicated a drastic reduction of α error rates.

	Semisolid		Thick Liquid		Thin Liquid	
	α	β	α	β	α	β
Exclusion A	22.2 %	6.4 %	7.3 %	4.0 %	9.2 %	17.2 %
Exclusion B	1.8 %	6.4 %	3.6 %	4.0 %	0.5 %	17.2 %

Table 4.12 Total tally and averages of α and β error rates of classification of the two normal and abnormal groups. Legend: α = the chance that a swallow was falsely classified as abnormal; β = the chance that a swallow was falsely classified as normal; Exclusion A – excluding subjects with one swallow only; Exclusion B – excluding eldest adult subjects (only applicable for normal subjects), and subjects contributing only one swallow.

4.5 Chapter Summary

This Chapter has presented the detailed results stemming from this research. Corroboration was reported for swallowing sounds and mechanics, as well as swallowing sounds and WDT boundaries. Of the 24 features extracted, eleven were found to be irreplaceable. Feature reduction also highlighted the importance of Group T features (Time Duration and Waveform Dimension). Significance was abundant within Magnitude and Power Groups; however scarce within Frequency and Normality Groups. Overall classification indicated high training and mediocre testing accuracies during the preliminary round; however, testing accuracies appreciably improved during the screening process. Overall classification error rates were markedly low, with the 50% threshold value. In the following Chapter, relationships and connections, both obvious and obscure, are discussed in terms of the resultant facts and figures.

CHAPTER 5

DISCUSSION

Time, frequency, audio and video domain analysis is required to fully identify, extract and compare swallowing sound signals. First, identification of individual swallowing sound signals depends upon the detection of IDS and FDS signals (both time and frequency domains). These sounds are consistently sharp, crackling and at times brilliant, as described by [Hamlet et al., 1994]. On average, IDS was 33 milliseconds in duration in comparison with 10 to 30 milliseconds reported in [Vice et al., 1990]. The difference may be due to different video digitization rate. The IDS signals are initiated by a prominent excursion with a rapid rise, graduating into activity of lower amplitude and lower frequency of repetition; whereas, the FDS signals are less prominent and more variable in amplitude and morphology (Figures 4.1 and 4.2). Physically (audio and video domains), the IDS correspond to the subsequent opening or relaxing of the crico-pharyngeus sphincter and shifting of the bolus into the pharynx (once the epiglottis gradually moves downward, which may take up to 100 msec). After the explosive shift, deglutition continues until the epiglottis is returned (which occurs at a much faster rate than its initial downward movement; i.e., is at most approximately 33 msec or 1 frame) and the airways are reopened. These final movements correspond to the FDS (although they are not always audible). Overall, consistent identification of IDS and FDS are congruent with the findings of both Bosma and his team [Vice et al., 1995] and Kahrilas and his co-workers (who studied the swallowing sound in 3D with two videofluoroscopic cameras and simultaneous MRI) [Kahrilas et al., 1995]. Furthermore, this identification

greatly assists in the extraction of individual swallows from series of breath and swallowing sound signals.

Second, swallowing sound signals are compared in terms of their characteristic segments and sections; i.e., comparison between manual timings and WDT boundaries or segments, along with Opening and Transmission sections. To begin with, WDT segmentation detects sound signal segments, such as 'click' (or IDS) and 'gulp,' within all swallowing sound signals (Figures 3.3. and 3.4). Manual recorded transcriptions (helped especially by video analysis) corroborate these timings, and thus relate characteristic segments found by WDT segmentation with actual swallowing mechanism stages.

With the successfulness of WDT as a signal segmentation tool, characteristic Opening and Transmission sections are observed and defined (again, transcription comparisons corroborate with section boundaries). Subsequently, features are calculated with various degrees of significance. In general, a high degree of overlap is observed in all 24 calculated features between normal and abnormal swallowing signal characteristics (i.e., large standard deviations are observed). This result is entirely expected, as the abnormal swallowing sound signals are actually marginally normal (or the successful, non-aspirated, swallows generated by neurologically impaired subjects). Thus, ensuing classification (to detect abnormal swallowing sound signals regardless of their marginal normalcy) must overcome this commonality factor; i.e., shared or overlapping characteristics observed between normal and abnormal groups.

From Opening and Transmission sections, and of the five major feature groups (T, M, F, P, N), Group T is the most important yielding the highest individual classification

accuracy. In addition and within Group T, Time Duration features are deemed most important (as they altogether yield higher classification accuracies than Waveform Dimension features). Within Time Duration, of Group T, the most notable feature is Total Time Duration. Although excluded from the reduced feature set, due to its strong correlation with Time Duration of the Opening and Transmission sections, Total Time Duration is a frequent feature calculated in deglutition studies. For our normal subjects, Total Time Duration feature values (averaged across ten subjects, ages 3-54) range from 790 to 880 msec (Table 4.1), differing slightly from total deglutition apnea duration values reported in [Kahrilas et al., 1995; Boiron et al., 1997; Klahn & Perlman, 1999; Perlman et al., 2000]. Kahrilas and his associates reported 870 msec from upper esophageal sphincter opening to airway reopening, with subsequent swallow in between, for one normal subject (male, age 35) [Kahrilas et al., 1995]. The remaining researchers reported apnea durations stemming from bolus viscosity and volume studies (with 92, 15 and 12 healthy normal adult participants, respectively). The averaged durations approximate 600 msec, 800 msec, and 750 msec [Boiron et al., 1997; Klahn & Perlman, 1999; Perlman et al., 2000]. Another notable result from Time Duration feature calculations is that all Time Duration features are significantly different for semisolid and thick liquid textures; i.e. abnormal swallowing sound signals are significantly longer than normal swallowing sound signals. This result is in accordance with those found in a recent pilot study comparing healthy controls and subjects with GERD (gastroesophageal reflux disease) [Mendell & Logemann, 2002]. For several pharyngeal swallow events, such as pharyngeal response and transit time, base of tongue movement to the posterior pharyngeal wall, hyoid movement and crico-pharyngeus sphincter opening, the GERD

group demonstrated longer time durations. In addition, all of the remaining recorded events, such as laryngeal elevation and closure, were also longer in duration in the GERD group, but not significantly so [Mendell & Logemann, 2002]. Further discussion follows regarding the speculative bases for this and other normal/abnormal feature differences.

Also within Group T, are Waveform Dimension features, which reveal significantly lower complexity in abnormal swallowing signals. This finding is interesting as WDT values may be potential indicators of specific deficiencies within the swallowing mechanism. When an aspiration event occurs, an error in the normal swallow sequence must have occurred. This error may be either a few milliseconds of delay of the airway protective features, a mistiming of the breaths and swallow or a poor clearance of the swallowed bolus, among others. The neurological impairments that underlie these problems in the swallowing mechanism are numerous (as mentioned in Section 2.2). Regardless of the cause of swallow difficulty, our data suggest that abnormal swallowing sounds lack a characteristic component compared to normal swallowing sounds, and that this is primarily reflected in waveform dimension values.

Following Group T, we find Group M features informative with Magnitude features yielding significant differences between normal and abnormal classes for all textures. Also, we note that discerning between consistencies through acoustical means is difficult, even with several Magnitude features yielding significant differences between bolus textures (e.g., Opening section Magnitude; thick liquid-thin liquid and thin-liquid-semisolid).

Of interest, are the Frequency and Power features of Group F and P. Specifically, we note an Fpeak of 500-900Hz, in comparison with the "wet" or water swallow,

reportedly reaching 400-600Hz and 1,000Hz (for the signal beginning and ending respectively) in [Mackowiak et al., 1967]. Also, averaged Fpeak, Fmax and Fmean are averaged at 800Hz, 2,700Hz, and 1,250Hz that can be compared to the 300Hz, 700Hz, and 350Hz reported in [Mussel & Miyamoto, 1992] for breath sound signals. This comparison indicates the dominating high frequency components of a swallow within the breath and swallowing sound signal. We observed significant differences between all Power features for semisolid data, with Figure 4.3 depicting the expected overlap. This illustration depicts the quintessential basis for our study; i.e. abnormal swallowing sound signals are distinct and yet parallel versions of normal swallowing sound signals.

The Normality features of Group N revealed Skewness and Kurtosis features indicating both normal and abnormal swallowing sound signals failed the normality test (i.e. swallowing sound signal segments). One recent study described 94 aspirated swallowing sound signals from children as non-stationary with departures from normality [Chau et al., 2002]; however, it was not clear what section of swallowing sound was studied. Given the fact that normally physicians halt the videofluoroscopy after observing one or two aspirations, it is not clear if their data included the swallows in which aspiration had occurred or the swallows of subjects at risk of aspiration. Nevertheless, their findings are congruent with the results of this study, in terms of swallowing sound signals (both normal and abnormal) failing the normality test.

Of the 24 features extracted, eleven were selected as the optimum feature set with the majority (all but one) yielding significant differences between normal and abnormal swallowing sound signals, for at least one bolus texture. The feature exception was Waveform Dimension of the Transmission section. We speculate that the extraneous

bolus transmission sounds ('gulp' and so on) frequently emanating from this section interfere with the true characteristic sound of the Transmission section.

With the optimum feature set, discriminant analysis proved successful as a classification tool. The classification was achieved in two stages. At stage one, discriminant analysis was used to classify each swallow to either normal or abnormal group. At the second stage however, screening the class of the majority of the swallows of a subject, determined the class of that subject. As a rule of thumb, we classified a subject as a subject at risk if more than 50% of his/her swallows were classified as abnormal and vice versa. However, the results showed that we could achieve the same accuracy at the second stage of screening with a threshold of 70 %. The secondary screening classification significantly improved all classification accuracies, for all food textures. The most important results stemmed from thick liquid, as this texture contained the largest number of swallowing sound signals. Thus, for thick liquid, final classification accuracies were 100.0% for both normal and abnormal subjects. For thin liquid, final classification accuracies were 90.0% (9 of 10 subjects) and 100% for normal and abnormal subjects, respectively. Of the ten normal subjects tested, one subject – the eldest normal subject (age 54) – was misclassified; i.e. he was classified as at risk of aspiration. As a result of this interesting finding, we postulate that age increases a normal subject's likelihood to experience some degree of swallowing dysphagia (for example, recurring aspiration). According to the error rates averaged in Table 4.12, the inclusion of two eldest adult subjects as controls contributed to higher misclassification percentages for all textures; i.e. the participating adults were classified as at risk or abnormal subjects. Without those two subjects, overall error rates were extremely low

(with the exception of thin liquid, β-rate). Overall, semisolid texture had approximately the lowest error rates, followed by thick and thin liquids (Exclusion B); however, semisolid was the least important texture utilized, as it was the safest to consume (i.e., the subjects aspirated less frequently with semisolid consistency than the other two textures). For semisolid, final classification accuracies were 77.8% (7 of 9 subjects) for normal subjects and 100.0% for abnormal subjects. Again, the two misclassified subjects were the two eldest control subjects (ages, 38 and 54). Of interest are the results of [Selley et al., 1990], where four groups of healthy subjects were tested to establish the normal behavior or swallowing mechanism. The groups, all without neurological diseases or swallowing problems, included 25 children (ages 2-11), 23 school pupils (ages 11-18), 15 university students and staff (ages 18-30) and 18 elderly adults (60-90). First, Selley and his associates note that the complex feeding respiratory pattern in a given adult is repeated in 95% of swallowing events. Second, their study reveals maturation of the feeding respiratory pattern in the teenage years and remarkable consistency thereafter [Selley et al., 1990]. However, Selley and associates focused mainly on respiratory patterns (i.e. percentage of expirations immediately following the swallowing sound signal) in addition to time duration. On the other hand are the speculations of Fillion and Kilcast, who conjecture that dysphagia, in the form of poor oral sensory integrity and consequently bolus formation and movement, stems from age-related changes in muscle strength and a generalized decrease in epithelial sensitivity (caused by a decrease in the number of nerve endings with increased age) [Fillion & Kilcast, 2001]. Moreover, in another recent study, Hiss and her associates test 60 adults (20 adults per age group; 20-39, 40-59 and 60-83 years old) and corroboratively report increased swallowing apnea

durations with age [Hiss et al., 2001]. In addition, oro-pharyngeal pressure duration and hypo-pharyngeal pressure in amplitude and duration reportedly increase with older individuals in comparison with young adults [Hiss et al., 2001]. Thus, in aging individuals, changes in swallowing sound signal are directly correlated to age-related changes in anatomy and physiology [Hiss et al., 2001]. This may explain the reason of our three older misclassified control subjects. Overall, discriminant analysis proved successful as a preliminary and secondary classification tool differentiating between the sound attributes of normal and abnormal swallowing sound signals.

CHAPTER 6

CONCLUSION AND RECOMMENDATIONS

6.1   Concluding Remarks

This thesis proposes a method to divide the swallowing sound signal into characteristic segments using WDT, and to investigate whether there are any differences between normal and abnormal (or marginally normal) swallowing sounds. To start, Chapter 2 reviewed the swallow (anatomy, pathology) and provided ample background for the systemic methodology of Chapter 3. Consequent results were reported in Chapter 4, with various implications and connections discussed in Chapter 5.

Overall, our initial two hypotheses were proven to be correct. First, a direct relationship was revealed between characteristic segments (IDS, and so on) with actual stages of swallowing mechanism. Corroboration of the swallowing sound signals in time and frequency domains, and with audio and video data reveal that both Opening and Transmission sections stem from physiological changes within the throat. In addition, the proposed method of adaptive segmentation, based on waveform dimension trajectory (WDT), has shown good ability to detect significant changes within the swallowing sound signal; specifically the 'clicks' (produced by IDS and/or FDS) and other bolus transition sounds (such as 'gulp'). Second, significant attribute differences were found between the Opening and Transmission sections of normal and abnormal swallowing sound signals. Within the Groups of Time Duration, Waveform Dimension, Magnitude, Frequency and Power, and 24 subsequently calculated features, eleven prominent features were utilized in a reduced feature set. With the exception of the Frequency features, each

Group had at least one feature remaining in the final reduced set. In terms of classification via discriminant analysis, the two-step diagnosis method resulted in high training and mediocre testing accuracies in the preliminary round, and high testing accuracies in the secondary round – with the notable exception of several adult control subjects.

6.2   Suggestions for Future Work

Although the proposed method is reliable in detecting normal and abnormal subjects, more information gathering, testing and refining is required. First, we used a convenient retrospective sample of subjects (both normal and abnormal) for testing the proposed method. It is imperative to collect more subjects, in homogeneous groups, in order to fully ascertain the capabilities of the proposed methods. Second, with rigorous data collection and an extended signal database, further testing is essential. In conjunction with clinical swallowing assessment, and fully updated clinical records, it is also desired to reduce the variation calculated per characteristic feature per group and to incorporate additional breath-swallow information into the classification model. In addition, classification may be expanded beyond the current normal/abnormal groupings; i.e., abnormal swallowing sound signals could be divided further into disease-specific groupings. Third, during signal analysis, timing and WDT boundaries were manually selected. A problem that requires additional study is the subjectivity of these selections and reduction of time spent on this preliminary constraint. In addition, extraneous sounds are currently included in WDT segmentation and subsequent analyses, which may affect the resultant characteristic features. Semi-automated boundary selection, and subsequent

extraction of excessive sounds, may provide temporary alleviation of these hindrances. On the whole, our future efforts will be directed toward expanding our database, vigorous testing and creating a semi-automated segmentation program.

Regardless of limitations, the proposed methods hold promise as a tool in the evaluation of swallowing dysfunction; and more importantly, in understanding the complexity of the swallowing process in individuals with and without neurological impairment Stemming from our results, long-term objectives include viable, non-invasive diagnosis of subjects (normal and abnormal) based on the analysis of externally recorded swallowing sound signals. The results of this research pave the way towards reducing the need for videofluoroscopic swallowing studies and assisting in the overall clinical assessment of swallowing sound signals.

REFERENCES

[Basmajian, 1970]

Basmajian, J.V., *Primary Anatomy*. Maryland: The Williams & Wilkins Company, 1970, Chapters 6 and 7, pp.211-249.

[Bishop, 1995]

Bishop, C.M., *Neural Networks for Pattern Recognition*. New York: Oxford University Press, 1995, Chapter 1, pp. 1-32.

[Boiron et al., 1997]

Boiron, M., Rouleau, P., Metman, E.H., "Exploration of Pharyngeal Swallowing by Audiosignal Recording," *Dysphagia*, vol. 12, pp. 86-92, 1997.

[Casas et al., 1995]

Casa, M.J., McPerson, K.A., Kenny, D.J., "Durational Aspects of Oral Swallowing Neurologically Normal Children and Children with Cerebral Palsy: An Ultrasound Investigation," *Dysphagia*, vol. 8, pp. 359-367, 1995.

[Chau et al., 2002]

Chau, T., Casas, M., Berall, G., Kenny, D., "Testing the Stationarity and Normality of Paediatric Aspiration Signals," *Proc. of IEEE EMBS '02 Conf.*, vol.1, pp. 186-187, 2002.

[Cichero & Murdoch, 1998]

Cichero, J.A.Y., Murdoch, B.E., "The Physiologic Cause of Swallowing Sounds: Answers from Heart Sounds and Vocal Tract Acoustics," *Dysphagia*, vol. 13, pp. 39-52, 1998.

[Cichero & Murdoch, 2002]

Cichero, J.A.Y., Murdoch, B.E., "Detection of Swallowing Sounds: Methodology Revisited," *Dysphagia*, vol. 17, pp. 40-49, 2002.

[Dempsey et al, 1990]

Dempsey, J.E., Vice, F.L., Bosma, J.F., "Combination of Cervical Auscultation and Videoradiography in Evaluation of Oral and Pharyngeal Dysphagia" (Abstract), *Symposium of Dysphagia*, Johns Hopkins Hospital, March 1990.

[Esteller et al., 2001]

Esteller, R., Vachtsevanos, G., Echauz, J., Litt, B., "A Comparison of Waveform Fractal Dimension Algorithms," *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, vol. 48, no.2, February 2001.

[Fillion & Kilcast, 2001]

Fillion, L., Kilcast, D., "Food Texture and Eating Difficulties in the Elderly," *Food Industry Journal*, vol. 4, no. 1, pp. 27-33, 2001.

[Firmin et al., 1997]

Firmin, H., Reilly, S., Fourcin, A., "Non-Invasive Monitoring of Reflexive Swallowing," *Speech, Hearing and Language*, vol. 10, 1997.

[Gewolb et al., 2000]

Gewolb, I.H., Bosma, J.F., Taciak, V.L., Vice, F.L., "Abnormal Developmental Patterns of Suck and Swallow Rhythms During Feeding in Preterm Infants with Bronchopulmonary Dysplasia", *Developmental Medicine and Child Neurology*, vol. 43, no. 7, pp. 454-459, July 2001.

[Gewolb et al., 2003]

Gewolb, I.H., Bosma, J.F., Reynolds, E.W., Vice. F.L., "Integration of Suck and Swallow Rhythms During Feeding in Preterm Infants with and without Bronchopulmonary Dysplasia," *Developmental Medicine and Child Neurology*, vol. 45, no. 5, pp. 344-348, May 2003.

[Hamlet et al., 1994]

Hamlet, S. Penney, D.G., Formolo, J., "Stethoscope Acoustics and Cervical Auscultation of Swallowing," *Dysphagia*, vol. 9, pp.63-68, 1994.

[Hiss et al., 2001]

Hiss, S.G., Treole, K., Stuart, A., "Effects of Age, Gender, Bolus Volume, and Trial on Swallowing Apnea Duration and Swallow/Respiratory Phase Relationships of Normal Adults," *Dysphagia*, vol. 16, pp. 128-135, 2001.

[Kahrilas et al, 1995]

Kahrilas, P.J., Lin, S., Chen, J., Logemann, J.A., "Three-dimensional Modeling of the Oropharynx During Swallowing," *Radiology*, vol. 194, no.2, pp.575-579, 1995.

[Katz, 1998]

Katz, M.J., "Fractals and the Analysis of Waveforms," *Computers in Biology and Medicine*, vol. 18, no. 3, pp. 145-156, 1998.

[Kinsner, 1994]

Kinsner, W., "Batch and Real-Time Computation of a Fractal Dimension Based on Variance of a Time Series," *Technical Report, DEL94-6*. Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada, 1994.

[Klahn & Perlman, 1999]

Klahn, M.S., Perlman, A.L., "Temporal and Durational Patterns Associating Respiration and Swallowing," *Dysphagia*, vol. 14, pp. 131-138, 1999.

[Last, 1978]

Last, R.J., *ANATOMY Regional and Applied.* London: Churchill Livingstone, 1978, Chapter 6, pp. 405-430.

[Lazareck & Moussavi, 2002]

L.J. Lazareck, Z. Moussavi, "Adaptive Swallowing Sound Segmentation By Variance Dimension," *Proc. EMBEC 02 European Medical and Biological Engineering Conference*, vol. 1, pp. 492-493, 2002.

[Lear et al., 1965]

Lear, C.S.C., Flanagan, J.B., Moorrees, C.F.A., "The Frequency of Deglutition in Man," *Arch Oral biology*, vol. 10, pp. 83-99, 1965.

[Leopold & Kagel, 1997 (a)]

Leopold, N.A., Kagel, M.C., "Dysphagia – Ingestion or Deglutition?: A Proposed Paradigm," *Dysphagia*, vol. 12, pp. 202-206, 1997.

[Leopold & Kagel, 1997 (b)]

Leopold, N.A., Kagel, M.C., "Pharyngo-Esophageal Dysphagia in Parkinsons' Disease," *Dysphagia*, vol. 12, pp. 11-18, 1997.

[Logemann, 1986]

Logemann, J.A., *Manual for the Videoflourographic Study of Swallowing, 2nd Edition*, Boston College-Hill Press, 1986.

[Logemann, 1995]

Logemann, J.A., "Dysphagia: Evaluation and Treatment," *Folia Phoniatrica et Logopedica*, vol. 47, no. 3, pp. 140-164.

[Mackowiak et al., 1967]

Mackowiak, R.C., Brenman, H.S., Friedman, M.H.F., "Acoustic Profile of Deglutition," *Proceedings Society for Experimental Biology and Medicine*, vol. 125, pp.1149-52, 1967.

[Matlab, 2000]

Matlab, The Language of Technical Computing, Computer Software Help Files, The Mathworks, Inc., version 6.0.0.88, 2000.

[Mendell & Logemann, 2002]

Mendell, D.A., Logemann, J.A., "A Retrospective Analysis of the Pharyngeal Swallow in Patients with a Clinical Diagnosis of GERD Compared with Normal Controls: A Pilot Study," *Dysphagia*, vol. 17, pp. 220-226, 2002.

[Mirrett et al., 1994]

Mirrett, P.L., Riski, J.E., Glascott, J., Johnson, V., "Videofluroscopic assessment of dysphagia in children with severe cerebral palsy," *Dysphagia*, vol. 8, pp. 209-214, 1993.

[Moussavi et al., 1998]

Moussavi, Z.K., Leopando, M.T., Rempel, G.R., "Automated Detection of Respiratory Phases by Acoustical Means," *Proc. of the 20th Annual Int. Conf., IEEE-EMBS*, vol. 20, no. 1, pp. 21-24, 1998.

[Mussel & Miyamoto, 1992]

Mussell, M.J., Miyamoto, Y., "Comparison of Normal Respiratory Sounds Recorded from the Chest and Trachea at Various Respiratory Air Flow Levels." *Frontiers Medicine and Biology Engineering*, vol. 4, no. 2, 1992, pp. 73-85.

[Norušis, 1994 (a)]

Norušis, M.J., *SPSS Professional Statistics 6.1*, SPSS Inc., Chicago, IL., 1994, Chapter 1, pp. 1-45.

[Norušis, 1994 (b)]

SPSS 6.1 Base System User's Guide, Part 1, UNIX Version, SPSS Inc., Chicago, IL., 1994.

[Norušis, 1994 (c)]

Norušis, M.J., *SPSS 6.1 Base System User's Guide, Part 2, UNIX Version*, SPSS Inc., Chicago, IL., 1994.

[Palisano et al., 1997]

Palisano, R., Rosenbaum, P., Walter, S., Russel, D., Wood, E., Galuppi, B., "Gross Motor Function Classification System for Cerebral Palsy," *Developmental Medicine Child Neurology*, vol. 39, pp. 214-223, 1997.

[Palmer et al., 1993]

Palmer, J.B., Kuhlemeier, K.V., Tippett, D.C., Lynch, C., "A Protocol for the Videofluorographic Swallowing Study," *Dysphagia*, vol. 8, pp. 209-214, 1993.

[Papoulis et al., 2002]

Papoulis, A., Pillai, S.U., *Probability, Random Variables and Stochastic Processes*. New York: McGraw Hill Companies, Inc., 2002.

[Peitgen, et al., 1992]

Peitgen, H-O, Jürgens, H., Saupe, D., *Chaos and Fractals New Frontiers of Science*. New York: Springer-Verlag, 1992, Chapter 9, pp. 457-503.

[Perlman et al., 2000]

Perlman, A.L., Ettema, S.L., Barkmeier, J., "Respiratory and Acoustic Signals Associated with Bolus Passage During Swallowing," *Dysphagia*, vol. 15, pp. 89-94, 2000.

[Qureshi et al., 2000]

Qureshi, M.A., Vice, F.L., Taciak, V.L., Bosma, J.F., Gewolb, I.H., "Changes in Rhythmic Suckle Feeding Patterns in Term Infants in the First Month of Life," *Developmental Medicine and Child Neurology*, vol. 44, no. 1, pp. 34-39, January 2002.

[Reddy et al., 1988]

Reddy, N.P., Canilang, E.P., Grotz, R.C., Rane, M.B., Casterline, J., Costarella, B.R., "Biomechanical Quantification for Assessment and Diagnosis of Dysphagia," *IEEE Engineering in Medicine and Biology Magazine*, vol. 7, no. 3, pp. 16-20, September 1988.

[Reddy et al., 1990]

Reddy, N.P., Costarella, B.R., Grotz, R.C., Canilang, E.P., "Biomechanical Measurements to Characterize the Oral Phase of Dysphagia," *IEEE Transactions on Biomedical Engineering*, vol. 37, no. 4, pp. 392-397, April 1990.

[Rempel et al., 1999]

Rempel, G., Pasterkamp, H. and Moussavi, Z., "The Effect of Altering the Consistency of Food on the Breath and Swallow Pattern of Young People with Cerebral Palsy," *53rd Annual Congress of American Academy for Cerebral Palsy and Development Medicine*, Sept. 1999.

[Selley et al., 1990]

Selley, W.G., Flack, F.C., Ellis, R.E., Brooks, W.A., "The Exeter Dysphagia Assessment Technique," *Dysphagia*, vol. 4, pp. 227-235, 1990.

[Selley et al., 1998]

Selley, W.G., Ellis, R.E., Flack, F.C., "The 'Cardiac Hypothesis' Explanation of Swallowing Sounds: A Revision and Alternative Interpretation of the Data," *Dysphagia*, vol. 13, pp. 232-234, 1998.

[Takahashi et al., 1994 (a)]

Takahashi, K., Groher, M.E., Michi, K., "Methodology for Detecting Swallowing Sounds," *Dysphagia*, vol. 9, pp. 54-62, 1994.

[Takahashi et al., 1994 (b)]

Takahashi, K., Groher, M.E., Michi, K., "Symmetry and Reproducibility of Swallowing Sounds," *Dysphagia*, vol. 9, pp. 168-173, 1994.

[Tarrant et al., 1997]

Tarrant, S.C., Ellis, R.E., Flack, F.C., Selley, W.G., "Comparative Review of Techniques for Recording Respiratory Events at Rest and During Deglutition," *Dysphagia*, vol. 12, pp. 24-38, 1997.

[Vice et al., 1990]

Vice. F.L., Heinz, J.M., Giuriati, G., Hood, M., Bosma, J.F., "Cervical Auscultation of Suckle Feeding in Newborn Infants," *Developmental Medicine and Child Neurology*, vol. 32, pp. 760-768, 1990.

[Vice et al., 1995]

Vice, F.L., Bamford, O., Heinz, J.M., Bosma, J.F., "Correlation of Cervical Auscultation with Physiological Recording During Suckle-Feeding in Newborn Infants," *Developmental Medicine and Child Neurology*, vol. 37, pp. 167-179, 1995.

[Webster's Dictionary, 1988]

Neufeldt, V., ed., *Webster's New World Dictionary of American English*. Cleveland & New York: Simon & Schuster, Inc., Third College Edition, 1988.