Nasometric Measurement and the Classification of Resonance Disorders: Equipment Evaluation and a Tentative Classification System

Gillian de Boer

Master of Science

Department of Speech-Language Pathology

University of Toronto

2013

Abstract

Resonance disorders due to cleft palate and other aetiologies are frequently assessed in conjunction with nasometry. The most commonly used instrument is the Nasometer by KayPentax. A new model Nasometer 6450 was compared to an older model 6200 using both synthetic and speech stimuli. There was a particular focus on test-retest variability of the instrument. The Nasometers were found to yield comparable results. The inter session test-retest variability ranged from six to eight points, depending on the stimulus. The Nasometer 6450 was then used to collect nasalance scores of simulated resonance disorders. A discriminant analysis was applied to these scores. The resultant formulas were moderately successful in predicting perceived resonance when applied to pre-existing data sets.
Acknowledgments

I would like to thank Dr. Tim Bressmann for his insight, support and guidance. I’m especially grateful for the times he lent me his learned glasses when mine were out of focus.

I would also like to thank the members of my supervisory committee, Dr. Pascal van Lieshout and Dr. Kiran Kulkarni for their direction and advice.

My friends and cohorts at the Voice and Resonance Lab, Susan Harper, Sina Koch, Bojana Radovanovic, Suzi Dumitrescu and all the volunteers, brought perspective and laughter to the table.

Finally, I’m forever grateful to my “hubby” Michael, for his love and encouragement.
# Table of Contents

Abstract ........................................................................................................................................... ii

Acknowledgments .......................................................................................................................... iii

List of Tables ................................................................................................................................... vi

List of Figures ............................................................................................................................... vii

1. Introduction ..................................................................................................................................1
   1.1 Normal Resonance .................................................................................................................2

   1.2 Disordered Resonance ............................................................................................................3
      1.2.1 Causes of Disordered Resonance ....................................................................................4

   1.3 Treatment of Hypernasal Resonance Disorders .................................................................5

   1.4 Diagnosis of Resonance Disorders .......................................................................................6
      1.4.1 Perceptual Assessment ....................................................................................................6

      1.4.2 Instrumental Assessment ................................................................................................7

   1.5 Study Objectives ....................................................................................................................8

2. Comparison of Nasalance Scores Obtained with the Nasometers 6200 and 6450.................10
   Abstract ......................................................................................................................................10

   2.1 Introduction ..........................................................................................................................11

   2.2 Study 1: Comparative Assessment of the Nasometers 6200 and 6450 with Square Wave
   Test Sounds ................................................................................................................................15
      2.2.1 Methods .........................................................................................................................15

      2.2.2 Results ...........................................................................................................................19

   2.3 Study 2: Comparative Assessment of the Nasometers 6200 and 6450 with Normal
   Participants ...............................................................................................................................22
      2.3.1 Participants ....................................................................................................................22

      2.3.2 Methods ........................................................................................................................23
List of Tables

Table 2.1  Nasalance scores for square wave test sounds in different stereo panoramas, recorded with the Nasometer 6200.

Table 2.2  Nasalance scores for square wave test sounds in different stereo panoramas, recorded with the Nasometer 6450.

Table 2.3  Mean nasalance scores and standard deviations by Nasometer model and session.

Table 2.4  Means and cumulative frequencies of test-retest differences by Nasometer model.

Table 3.1  Mean nasalance scores and nasalance distance with standard deviations for the Zoo Passage and the Nasal Sentences as measured with the Nasometer 6200 from Bressmann et al. (2006)

Table 3.2  Mean nasalance scores and nasalance distance with standard deviations for the Zoo sentence (oral) and the first Nasal sentence as measured with the Nasometer 6450 by condition (normal, hyponasal low, hyponasal high, hypernasal, mixed low and mixed high).

Table 3.3  Canonical discriminant function coefficients derived from two predictors (oral and nasal stimuli) and six simulated conditions (normal, hyponasal low, hyponasal high, hypernasal, mixed low and mixed high).

Table 3.4  Function values of group centroids for six simulated conditions (normal, hyponasal low, hyponasal high, hypernasal, mixed low and mixed high).

Table 3.5  Canonical discriminant function coefficients derived from two predictors (oral and nasal stimuli) and four simulated conditions (normal, hyponasal low, hypernasal and mixed low).

Table 3.6  Function values of group centroids for four simulated conditions (normal, hyponasal low, hypernasal and mixed low).

Table 3.7  Results for sensitivity and specificity of the discriminant functions when applied to the data set by Bressmann et al. (2006).
List of Figures

Figure 2.1 Measuring set-up for square wave test sounds.

Figure 2.2 Sound pressure level in dB (C) by frequency for the left and right loudspeaker.

Figure 3.1 Scatterplot of function values with group centroids

Figure 4.1 Hypothetical resonance matrix for normal, hyponasal, hypernasal and mixed nasality based on scores from the Zoo Passage (oral) and the Nasal Sentences using the Nasometer 6450.
1. Introduction

Speech articulation and resonance are produced in the vocal tract. They require the regulation of breath, the propulsion of air through the vocal folds, the separation and coupling of air between the oral and nasal cavities and the movement of the jaw, tongue and lips. With a sufficient transglottal pressure difference, the vocal folds vibrate to produce a specific source signal. This source signal is modulated by aerodynamic changes introduced by the pharynx and the articulating organs. If the velopharyngeal port is open, sound will travel to the nasal cavities. The coupling and de-coupling of sound between the oral and nasal cavities is referred to as the oral-nasal balance (Hixon, Hoit & Weismer, 2007; Peterson-Falzone, Hardin-Jones & Karnell, 2001). The acoustic measurement of this oral-nasal balance is of interest in the present study.

Oral-nasal balance is normal when sound resonates in the mouth for oral sounds and in the nose for nasal sounds such as m, n and ng. When either or both of these resonance chambers are compromised, the result is a resonance disorder. The impact on the individual can range from the temporary hyponasality of the common cold to the severe hypernasality that can develop in children born with a cleft palate (Kummer, 2008). Resonance disorders are typically diagnosed by speech-language pathologists who then recommend an appropriate treatment, be it surgery, a prosthetic device or speech therapy. The primary diagnostic tool is the clinician’s trained ear. The auditory-perceptual diagnosis is supplemented with instrumental diagnostics, which can include videofluoroscopy, nasoendoscopy and nasometry (Kuehn & Moller, 2000). Each of these tools has its benefits and drawbacks. The present study will focus on nasometry, more specifically the Nasometer by KayPentax which is internationally the most popular and most commonly used device (Kummer, 2008). The research is divided into two parts. In the first part of the study,
described in chapter 2, we compared the latest Nasometer model 6450 to an older Nasometer model 6200 and evaluated their respective test-retest variabilities. In the second part of the study, described in chapter 3, the Nasometer 6450 was used to assess simulations of disordered resonance, produced by normal speakers. Based on the simulations, a diagnostic scheme for a tentative classification of normal, hyponasal, hypernasal and mixed nasal resonance was developed.

1.1 Normal Resonance

The oral and nasal resonating chambers are separated by the hard and soft palate. The soft palate together with the upper lateral and posterior pharyngeal walls constitutes the velopharyngeal mechanism (VPM) or velopharyngeal sphincter. The VPM serves as a valve that regulates the opening and closing of the velopharyngeal port (Peterson-Falzone et al., 2001). It is a complex structure with multiple paired muscles which are involved in speech, swallowing, breathing, and middle ear pressure equalization. The VPM muscles involved in speech are the levator veli palatini (elevates the velum), the palatoglossus (lowers the velum and, when in a fixed position, contributes to elevation of the tongue dorsum), the superior constrictor (contracts the walls of pharynx), while the musculi uvulae contracts the velum along its length and forms the velar eminence. According to Kummer (2008), the role of the palatopharyngeus is not well understood but it is thought to contribute to the narrowing of the velopharyngeal port. Finally, the tensor veli palatine opens the Eustachian tube (Kummer, 2008)

Except for the tensor veli palatine, which is innervated by the trigeminal nerve, these muscles are innervated by the plexus pharyngeus, a network of fibers from the glossopharyngeal (IX) and vagus (X) cranial nerves (Kummer, 2008). Some fibres of the accessory (XI) and facial (VII)
nerves may also be involved (Peterson-Falzone et al., 2001). The sensory innervation is thought
to be supplemented by the lesser palatine branch of the trigeminal (V) nerve (Kummer, 2008).
The density of the sensory neurons has been reported to decrease from the oral cavity to the
pharynx (Grossman & Hattis, 1964; Kanagasuntheram, Wong & Chan, 1964; Kuehn & Perry,
2009), which may explain the limited proprioception speakers have of their VPM (Peterson-
Falzone, Trost-Cardamone, Karnell, & Hardin-Jones, 2006).

VPM closure patterns are individually varied. While about half of speakers close the
velopharyngeal port by elevating the velum and stretching it to reach the posterior wall (coronal
closure), others approximate the lateral walls only (sagittal closure). Many speakers combine the
sagittal and coronal closure patterns into a circular closure pattern. A variation of the circular
closure pattern is characterized by a hypertrophic eminence on the posterior pharyngeal wall
(circular closure with Passavant’s ridge) (Croft, Shprintzen & Rakoff, 1981; Kummer, 2008;
Peterson-Falzone et al., 2001).

1.2 Disordered Resonance

Resonance will be disordered if, the VPM does not effectively separate the oral cavity from the
nasal cavities, the VPM does not open and close when required, or if the sound cannot escape its
resonating chamber (Kummer, 2011; Peterson-Falzone et al., 2001, 2006). When excess sound
resonates in, and escapes from, the nasal cavities, the result is hypernasal speech. A blockage in
the pharynx or the nasal cavities will prevent the production of nasal consonants and causes
hyponasal speech (Kummer, 2011). Cul-de-sac resonance is often considered a form of
hyponasality where the sound is trapped in the nasal cavities or upper nasalpharynx, producing a
muffled quality (Kummer, 2008; Peterson-Falzone et al., 2001, 2006). When both hypernasality
and hyponasality are present, the resonance disorder is known as mixed nasality (Kummer, 2008; Peterson-Falzone et al., 2001).

### 1.2.1 Causes of Disordered Resonance

Structural causes of hypernasality include cleft palate, oronasal fistulae, congenitally short velums and lesions from oral cancers or traumatic injury. Hypernasality is also a common feature of dysarthria (Kummer 2008, Peterson-Falzone et al. 2001) whereby the neuromotor control of the VPM is compromised. Hypernasality can also be due to mislearning and is often heard in the speech of individuals with severe hearing impairment (Kim, Yoon, Kim, Nam, Park, & Hong, 2012). Chronic hyponasality is almost always due to a blockage (Kummer, 2011). Posteriorly, this could be hypertrophic tonsils, or adenoids, choanal stenosis (Kummer, 2011), a pharyngeal flap (D’Antonio & Scherer, 2009) or a prosthetic speech appliance (Karnell, Hansen, Hardy, Lavelle, & Markt, 2004), preventing air and sound from reaching the nasal cavities.

Anteriorly, hypertrophic turbinates, a deviated septum, stenotic nares, and maxillary retrusion would only allow some of the air and sound that reached the nasal cavities to project to the listener (Kummer, 2011). Among the hearing impaired, the lack of auditory feedback can also produce hyponasal speech, because the speakers overcompensate by constantly closing the VPM during speech (Kim et al., 2012). When no sound escapes the nasal cavities and the blockage is complete, it is termed denasality (Kummer, 2008; Peterson-Falzone et al., 2006). Among the cleft lip and palate population, cul-de-sac resonance can be due to a cleft-related deviated septum or stenotic nares (Kummer, 2008; Peterson-Falzone et al., 2006). However, in the general population, enlarged tonsils are a more common aetiology (Kummer, 2011). For mixed nasality, the patient will have a combination of causes from those listed for hypernasality and hyponasality. A typical scenario is that of a speaker with a unilateral cleft lip and palate who has...
both dysfunction of the VPM (causing hypernasality) and a deviated septum (causing hyponasality) (Kummer, 2008; Peterson-Falzone et al., 2001)

1.3 Treatment of Hypernasal Resonance Disorders

Hypernasal resonance disorders can be treated with speech therapy, with surgery and with prosthodontics devices. If it can be demonstrated that a hypernasal speaker can achieve closure of the VPM, then treatment will consist of speech therapy. As proprioception of the mechanism is low, the therapy usually involves some form of biofeedback, such as a mirror or an air paddle below the nares, a stethoscope, or a SeeScape (Kummer, 2008; Sell & Grunwell, 2001). Some may also benefit from continuous airway pressure (CPAP), a treatment for snoring and obstructive sleep apnea. The air pressure can be used in speaking exercises to help strengthen the velopharyngeal sphincter (Kuehn, 1991; Kuehn et al., 2002). The surgical interventions for hypernasality were developed for children with cleft lip and palate. After the initial palatal repair, which takes place between six and eighteen months (Watson, 2001), about 20% of patients will continue to have some degree of hypernasality (Kummer, 2008). Specific surgeries for the remediation of hypernasality typically involve a partial blockage of the velopharyngeal port (Witt, 2009). This is accomplished by lengthening the palate with an intravelar veloplasty or double opposing Z-plasty or by reducing the size of the opening of the nasopharynx with a pharyngeal flap or sphincter pharyngoplasty (Witt, 2009; Watson 2001).

Two major types of speech prostheses are speech-bulbs and palatal lifts. A speech-bulb has a bulb extension at the pharyngeal end which sits inside the velopharyngeal opening and reduces the escape of air and acoustic energy into the nasal cavities. It is used where there is a structural insufficiency, such as a cleft palate or defect related to oral cancer (Kummer, 2008). In patients
with a sufficiently long but neurologically incompetent velum, the palatal lift appliance lifts the velum. It is used when there is enough mass of tissue to close the velum, but inadequate coordination of the muscles (dysarthria, apraxia) to achieve closure (Kummer, 2008). Although they are successful in normalising resonance, prostheses can be uncomfortable to wear (Kummer, 2008). If the discomfort in the fitting or wearing of the appliance is not acceptable, an exotic alternative treatment option is a nasal obturator. Two recently described types included a one-way valve and thus allowing for inhalation, but not exhalation (Beukelman, Fager, Green, Hakel, & Marshall 2004; Suwaki, Nanba, Ito, Kumakura & Minagi, 2008). Although these devices produce cul-de-sac resonance for the speaker, they can modestly increase the intelligibility (Suwaki et al., 2008).

1.4 Diagnosis of Resonance Disorders

1.4.1 Perceptual Assessment

The primary method in the assessment of resonance disorders is the clinician’s trained ear (Kuehn & Moller, 2000). Listening to the patient’s speech in spontaneous speech and repeating standardised sentences and words is essential to determine the presence and type of resonance disorders. Some causes of resonance disorders can be found with a simple oral-pharyngeal exam. An opened mouth can reveal an oro-nasal fistula, enlarged tonsils, a congenitally short soft palate and a submucous cleft of the hard and/or soft palate. However, the function of the VPM during speech is concealed from view. There are fundamental issues with the reliability of perceptual judgments (Whitehill & Lee, 2008). These will be discussed further in chapter 3.
1.4.2. Instrumental Assessment

When a resonance disorder is detected, the three most commonly used instruments in cleft centres are videofluoroscopy, nasoendoscopy and nasometry (Kuehn & Moller, 2000). Videofluoroscopy and nasopharyngoscopy are direct assessment techniques, which allow the clinician to visualize the velopharyngeal mechanism during speech, while nasometry is an indirect acoustic signal based measure of the oral-nasal balance.

Multiview videofluoroscopy provides a radiographic assessment of velopharyngeal movement during speech production. It is sometimes supplemented with Barium for improved visualisation (Hinton, 2009). The three most used views are the lateral view, for anterior-posterior movement and visualizing the structure of the velum and posterior pharyngeal wall, the frontal view for lateral pharyngeal wall movement, and, the Townes or base view for visualisation of the velopharyngeal closure pattern. The images are interpreted together to give a complete picture of the structure and function (Moon, 2009). Measurements can be made to calculate the ratio of velopharyngeal closure and the size of the gap. Cooperation is usually not an issue, but due the exposure of radiation, the filming is usually limited to two minutes (Hinton, 2009).

With nasoendoscopy, a topical anaesthetic is applied to the nostril and a flexible fiberoptic endoscope is inserted to view the VPM from above. This allows the pattern and extent of closure to be determined. This is particularly helpful in the visualization of an occult submucous cleft (Hinton, 2009). Without the concern of radiation, the procedure can be used for longer periods of time and more frequently (Whitehill & Kim, 2008). However, cooperation from younger children is often problematic (Kuehn & Moller, 2000; Moon, 2009). Where possible, Mercer & Pigott
(2001) advocate the subsequent, or even simultaneous, recording of endoscopic and videofluoroscopy images because both methods complement each other.

There has been research on the acoustics of nasal resonance (Kataoka, Warren, Zajac, Mayo & Lutz, 2001; Lee, Cioeoa & Whitehill, 2004) but it has not led to clinically useful diagnostic procedures. As an alternative, nasometry is used to provide a quantitative acoustic assessment of oral-nasal balance. Although applied to other populations, nasometry’s primary role is the assessment of hypernasality in patients with cleft lip and palate (Kummer, 2008). It consists of two microphones mounted to the topside and underside of a metal sound separator plate placed between the nose and the upper lip. The accompanying software then computes a nasalance score based on the ratio of oral to nasal sound pressure levels. Higher scores are associated with hypernasality and lower scores with hyponasality (Dalston, Warren, & Dalston, 1991a, 1991b; Hardin, Van Denmark, Morris, & Payne, 1992; Kummer, 2008). As discussed in more detail in Chapter 3, the scores correspond relatively well with listeners impressions of disordered resonance. The scores from standardized passages are used to provide a quantitative measure of tracking within subject changes over time (Peterson-Falzone et al., 2001) and assist in assessment and treatment decisions (Kummer, 2008). While non-invasive, the VPM is not seen and its function can only be inferred.

1.5 Study Objectives

The goal of the present study was to develop a tentative assessment protocol based on nasometry values that can help a clinician differentiate between different resonance disorders. The research of this thesis was carried out in two stages, the first serving as a basis for the second.
In the first study, the Nasometer 6450 was characterized based on synthetic stimuli and human participants. The Nasometer 6450 was used in both studies in this thesis. It was introduced relatively recently in 2009. However, there are no published reports on the characteristics of the Nasometer II 6450 to date. Before setting out to develop the tentative classification system, the performance of the Nasometer 6450 needed to be compared to that of its predecessor. Establishing the performance of the Nasometer 6450 in comparison to the Nasometer 6200 and the variability of nasalance scores using a variety of speech and synthetic stimuli was the focus of Chapter 2.

In the second study, simulations of hyponasal, hypernasal and mixed resonance were used to create a dataset of disordered resonance nasalance scores. While normal resonance and hypernasality have been studied in depth, very little is known of the nasalance scores associated with hyponasality and mixed nasality. We propose to use two simple nasometry values as the basis for a statistical classification of resonance type based on linear discriminant analysis. The resulting diagnostic algorithm to distinguish between normal, hyponasal, hypernasal and mixed resonance was applied to two pre-existing data sets. One dataset comprised normal speakers and speakers with cleft palate. The second dataset comprised individuals with velopharyngeal dysfunction who had been fitted with palatal lift appliances.
2. Comparison of Nasalance Scores Obtained with the Nasometers 6200 and 6450

Abstract

Objective: The study had the goal of characterizing the new Nasometer 6450 in comparison to the older model 6200 using both synthetic test sounds and control participants. A particular focus of the investigation was on the test-retest variability of the instruments.

Materials and methods: The Nasometers 6200 and 6450 were compared using square wave test sounds. Six repeated measurements of oral, balanced and nasal test stimuli were recorded from 25 healthy female participants over an average of 35 days.

Results: The synthetic test sounds demonstrated that the two Nasometers obtained similar results for a range of frequencies. The results for the participants revealed that nasalance scores from the two instruments were within 1-2 points, depending on the test sentence. For both systems, variability in scores increased with the proportion of nasal consonants in the sentence. Test-retest variability was between 6-8 points for over 90% of the participants. Participants with higher nasalance scores for oral stimuli had higher between-session variability.

Conclusions: The Nasometers 6200 and 6450 can be expected to yield comparable results in clinical practice. Depending on the phonetic content of the test materials, clinicians should allow for 6-8 point between-session variability in either direction when interpreting nasalance scores.

---

1 An earlier version of this chapter was submitted to the Cleft Palate-Craniofacial Journal August 16, 2012, and accepted Jan 14, 2013
2.1 Introduction

Computerized nasometry is commonly used in clinical practice to supplement the clinician’s perceptual assessment of a client’s resonance disorder. It is primarily used for the assessment of hypernasality in patients with cleft lip and palate (Kummer, 2008). The nasalance score expresses the relative contribution of oral and nasal sound pressure levels to the patient’s speech and is calculated according to the formula: nasalance = nasal/(nasal+oral) • 100 (Fletcher, 1976). In comparison to established normative data, higher nasalance scores are associated with hypernasality, and lower scores are associated with hyponasality (Kummer, 2008; Dalston et al., 1991a, 1991b). Audible nasal emissions or nasal turbulence may also affect the magnitude of nasalance scores (Dalston et al., 1991b; Karnell, 1995). Nasalance scores can be used to supplement the clinical assessment and to quantify treatment outcomes (Kummer, 2008).

Three nasometry instruments have been developed, the NasalView (Tiger Electronics, Seattle WA), the OroNasal System (Glottal Enterprises Inc., Syracuse NY) and the Nasometer by KayPENTAX (KayPENTAX, Montvale, NJ) (Bressmann, 2005; Kummer 2008). The Nasometer is the most commonly used of these instruments for the measurement of nasalance. Different generations of the Nasometer have used different methods of signal processing. The original Nasometer 6200, first introduced in 1986, converted the sound pressure levels to direct current and sent calculations to the computer. It was not possible to record an audiofile with the Nasometer 6200. A special innovation of the Nasometer 6200 was that it used a band pass filter with a centre frequency of 500 Hz and a range of 300 Hz to emphasize the extra resonances in this frequency range that are associated with hypernasal speech (Fletcher & Bishop, 1970). Baken & Orlikoff (2000) have argued that this filter range was chosen a priori without a sound rationale.
Awan (1998) has pointed out that the signal filtering affects the researchers’ ability to do acoustic analyses on nasalance recordings. Notwithstanding these criticisms, subsequent Nasometer models have employed the same filter characteristics in their signal processing. The second generation Nasometer II 6400, first distributed in 2002, used a preamplifier together with a dedicated computer sound card. The latest model, the Nasometer II 6450, was introduced in 2009. It uses an external universal serial bus sound card and transfers a digital sound file to the computer. Nasalance scores reflect the average of the ratio (above 0) computed every 8 milliseconds.

Despite the differences in signal acquisition, the manufacturers maintain that the models should score within two nasalance points of each other. Studies comparing the first two nasometer models have been divided about this claim. Watterson, Lewis & Brancamp (2005) found that the Nasometer 6400 scored about one point higher than the Nasometer 6200 for oral and balanced stimuli. However, Awan, Omlar, & Watts (2011) found the Nasometer 6400 scored four to six points lower than the Nasometer 6200 for five oral vowel-loaded sentences. Since the Nasometer is an important clinical tool, it was a goal of this study to compare the measurements from the new Nasometer 6450 to the reference instrument, the historical Nasometer 6200. It is important for researchers and clinicians to understand how nasalance scores from the two nasometers can be compared.

Another goal of the current study was to assess the test-retest variability for the new Nasometer 6450 in comparison to the Nasometer 6200. The initial studies of immediate test-retest variability in nasalance scores of normal participants, measured with the Nasometer 6200, found that between 94-100% of recordings were within 3 nasalance points for three repeated recordings.
(Seaver, Dalston, Leeper, & Adams, 1991; Litzaw & Dalston 1992). Using oral and balanced sentences, variability among hypernasal participants was found to be greater (Watterson & Lewis, 2006). The immediate test-retest nasalance score variability of the Nasometer 6200 was within 5 points for 88% of the participants and within 9 points for 89% of participants when the headset and microphone were removed and replaced between readings. While this suggests that, in hypernasal participants, scores within 10 nasalance points may be within their normal range of variability, the authors noted that 61% of the participants were well within 5 nasalance points (Watterson et al., 2006).

In a study of nasalance testing materials for normal female speakers of Cantonese, Whitehill (2001) measured day-to-day variability. For oral stimuli (a sentence and paragraph), 95% of the returning participants were within 5 points of their previous score. For balanced stimuli, 93% of participants were within 8 points and for nasal stimuli, 96% of participants were within 10 points. Lewis, Watterson & Blanton (2008) compared immediate and long-term test-retest variability for normal participants using the Nasometer 6400. Long-term variability was greater but did not increase with the time between measurements. Re-testing was done on the same day, twice a day for five consecutive days, and once a week for three weeks. The stimuli consisted of the Turtle (oral) and the Mouse (balanced) passages. The short-term variability was less than 5 nasalance points for over 90% of the participants. However, the long-term variability of 90% of the participants was within a range of 6 to 8 nasalance points. This is an important observation because a clinician would be inclined to attribute meaning to an 8 point increase (worsening hypernasality) or a decrease (improvement).
These studies provide substantial information about the variability of nasalance scores obtained with the Nasometers 6200 and 6400. However, with the introduction of the new Nasometer 6450, it becomes important to assess the test-retest characteristics for this instrument and to compare it to the original Nasometer 6200. Also, knowledge of normal variation of nasalance scores obtained from nasal stimuli over time is limited to the study by Whitehill (2001) who demonstrated that nasal test stimuli had higher test-retest variability. In a similar vein, Watterson et al. (2006) found that hypernasal speakers had higher test-retest variability. This suggests that there could be a relationship between subjects’ nasalance scores and their mean test-retest variability.

Zajac, Lutz & Mayo (1996) investigated the relationship between the sensitivity of the microphones used in the Nasometer 6200. The manufacturer had not yet standardized the microphones issued with the Nasometer 6200 at the time. The authors found that the microphone characteristics influenced the nasalance scores. Other than this study, almost all of the research concerning nasometer measurements has been carried out with human participants. Synthetic test sounds have the advantage that they do not have the variability of human speakers or the variance introduced by setting up the equipment for each speaker. We therefore argue that it would also be desirable to characterize the new Nasometer 6450 in more detail using synthetic test sounds.

Based on the foregoing, the present study had as a first goal to investigate the Nasometer equipment using a set of synthetic test sounds to assess whether both instruments deliver stable measurements over a range of frequencies. The second goal was to investigate the differences in nasalance scores, if any, obtained with the Nasometer 6200 and the Nasometer 6450 using normal speaking participants. The third goal was to obtain the test-retest variability over time for
the Nasometer 6450 in comparison to the Nasometer 6200. Accordingly, the first null hypothesis for the study to be refuted was that the Nasometers 6200 and 6450 would yield identical results for all measures, synthetic or from human participants. The second hypothesis to be refuted was that the test-retest variability for all participants would yield identical results at all time points.

2.2 Study 1: Comparative Assessment of the Nasometers 6200 and 6450 with Square Wave Test Sounds

2.2.1 Methods

Nineteen square wave sound files of two second duration were created using the MultiSpeech 3700 software (KayPENTAX, Englewood Shores, NJ). The sounds were created with frequencies ranging from 105 to 305 Hz in 25 Hz increments. This was based on research by Zajac et al. (1996) who argued that this frequency range encompassed common human fundamental frequencies.

We also created a set of square waves with higher frequencies ranging from 355 to 805 Hz in 50 Hz increments. These test sounds were used to evaluate how the two Nasometers behaved in the band pass-filtered frequency range around the 500 Hz centre frequency. The 755 Hz and 805 Hz sounds were included to straddle the upper border of the Nasometer’s frequency range and to assess whether the filter cut-off affects the measurements. All square wave files were saved to *.wav format.

The Nasometer was calibrated using a single loudspeaker or sound source. The previous research by Zajac et al. (1996) also used a single loudspeaker to assess different Nasometer headsets. Here
a stereo configuration with two loudspeakers addressing the two microphones on the Nasometer headset was used. This was done to estimate the crossover between the two microphones and to assess the robustness of the nasometer measurements. The Goldwave digital audio editor (GoldWave Inc., St. John's, NL) was used to create five different stereo panoramas for each sound file: hard left, three-quarter left, balanced, three-quarter right, and hard right. The panoramas were chosen to reflect the increase in nasalance scores as more acoustic energy was directed towards the nasal microphone. When the left loudspeaker aligns with the oral microphone and the right loudspeaker aligns with the right speaker, this should, in theory, lead to measurements approximating 0%, 25%, 50%, 75% and 100% nasalance respectively. Due to acoustic spill over around the 25 dB separation plate, the real measures cannot be expected to be so clear cut (KayPENTAX, 2010). Nevertheless, it was expected that the nasalance scores for the hard left and three-quarter left recordings would be exact mirror images of the hard right and three-quarter right recordings.

The sound files were played from a Hewlett Packard Pavilion laptop with the headphone output volume set to “2”. The loudspeakers used were iHome iHM78 mini-speakers (iHome, Rahway, NJ). These loudspeakers are cylindrical in shape and angled at the top. To emphasize lower frequencies, a bass cabinet was extended from each loudspeaker’s base. With the extended bass cabinet, the lowest frequency that the loudspeakers could produce was 80 Hz. The frequency range of the loudspeakers was assessed to verify that the 105 Hz frequency could be produced by the speakers.

The sound energy emitted by the loudspeakers was measured with a Brul and Kjaer analog sound level meter model 2209 (Brul and Kjaer, Naerum Denmark). In keeping with the methods
used by Zajac et al. (1996), the weighting filter was set to C. This setting also has the advantage
that it accentuates loudness differences in the low frequencies. The time weighting was set to fast.
The tip of the sound level meter was 7 cm from the centre of either loudspeaker cone. The sound
pressure level for each frequency was recorded for the left and right loudspeakers separately.

The loudspeakers and the Nasometer headset were placed in a custom-made holder (Figure 2.1).
The holder was made from blocks of packing foam. There was a 1 cm foam spacer between the
loudspeaker and the Nasometer separator plate on either side. The loudspeaker cones were 2 cm
beneath the Nasometer microphones. In the default setup, the nasal microphone recorded from
the right loudspeaker and the oral microphone from the left loudspeaker. To control for
differences in the sound energy emitted from the loudspeakers themselves, the orientation of the
setup was switched for a part of the recordings so that the nasal microphone recorded from the
left loudspeaker and the oral microphone from the right loudspeaker.

Three repeated nasalance recordings were made for each frequency in the hard left, three-quarter
left, balanced, three-quarter right, and hard right stereo panoramas. With the Nasometer
separation plate and the microphones reversed, the test sounds in the three-quarter right, three-
quarter left and balanced panoramas were repeated. Nasalance scores were recorded three times
for each of the nineteen frequencies, eight sound panoramas and both Nasometer models,
resulting in a total of 912 nasalance scores. The Nasometer 6200 provides nasalance scores to
two decimal places while the Nasometer 6450 provides only integers. Before mean values and
standard deviations were calculated, the nasalance scores of the Nasometer 6200 were rounded to
the nearest integer using the rounding function in SPSS. The same operation was performed by
Watterson et al. (2005) when they compared the Nasometers 6200 and 6400.
Both Nasometers were used with their respective headsets. Before the recordings, the Nasometers 6200 and 6450 were calibrated with a pulse signal from the Nasometer 6450. The calibration readings were 49.65% for the Nasometer 6200 and 0.96 for the Nasometer 6450. The first calibration is computed as a “nasalance score” and the second is a ratio of the sound energy reaching the microphones. Both calibration scores were within the manufacturer’s recommended range.
2.2.2 Results

The results for the sound pressure level measurements of the two loudspeakers are displayed in Figure 2. The results indicate that the left loudspeaker was louder by an average of 1.78 dB (C) (SD 0.645) for all frequencies, which was statistically significant in a paired t-test (p < .01).

Figure 2.2 Sound pressure level in dB (C) by frequency for the left and right loudspeaker.

The mean nasalance values for the results for all frequencies tested are displayed in Tables 2.1 and 2.2. A first inspection of the results drew our attention to the mean nasalance scores that were obtained for the three-quarter right and three-quarter left sound files for the 755 Hz square wave, measured with the Nasometer 6200. Statistical analysis confirmed that these scores were outliers. The same sound file did not present different results for the Nasometer 6450.
Table 2.1 Nasalance scores for square wave test sounds in different stereo panoramas, recorded with the Nasometer 6200.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Left Mean</th>
<th>SD</th>
<th>¼ Left Mean</th>
<th>SD</th>
<th>Balanced Mean</th>
<th>SD</th>
<th>¾ Right Mean</th>
<th>SD</th>
<th>Right Mean</th>
<th>SD</th>
<th>Rev ¼ Left Mean</th>
<th>SD</th>
<th>Rev Balanced Mean</th>
<th>SD</th>
<th>Rev ¾ Right Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>805</td>
<td>11 0</td>
<td>25</td>
<td>0</td>
<td>47</td>
<td>0</td>
<td>72.33</td>
<td>0.58</td>
<td>88</td>
<td>0</td>
<td>76</td>
<td>0</td>
<td>54</td>
<td>0</td>
<td>29</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>755</td>
<td>7.33 0.58</td>
<td>42</td>
<td>0</td>
<td>47</td>
<td>0</td>
<td>60.67</td>
<td>0.58</td>
<td>91.67</td>
<td>2.31</td>
<td>60</td>
<td>0</td>
<td>54.67</td>
<td>0.58</td>
<td>41</td>
<td>0</td>
<td>41</td>
</tr>
<tr>
<td>705</td>
<td>9 0</td>
<td>23</td>
<td>0</td>
<td>47</td>
<td>0</td>
<td>72</td>
<td>0</td>
<td>90</td>
<td>0</td>
<td>79</td>
<td>0</td>
<td>56</td>
<td>0</td>
<td>30.67</td>
<td>0.58</td>
<td>30.67</td>
</tr>
<tr>
<td>655</td>
<td>9 0</td>
<td>24</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>71</td>
<td>0</td>
<td>90</td>
<td>0</td>
<td>78</td>
<td>0</td>
<td>55.33</td>
<td>0.58</td>
<td>31</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>605</td>
<td>7.33 0.58</td>
<td>25</td>
<td>0</td>
<td>46.33</td>
<td>1.15</td>
<td>71</td>
<td>0</td>
<td>89</td>
<td>0</td>
<td>78</td>
<td>0</td>
<td>55</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>555</td>
<td>10 0</td>
<td>22.33</td>
<td>0.58</td>
<td>47</td>
<td>0</td>
<td>74</td>
<td>0</td>
<td>89</td>
<td>0</td>
<td>80.67</td>
<td>0.58</td>
<td>56</td>
<td>1.73</td>
<td>29.33</td>
<td>0.58</td>
<td>29.33</td>
</tr>
<tr>
<td>505</td>
<td>8 0</td>
<td>18</td>
<td>0</td>
<td>45</td>
<td>0</td>
<td>74</td>
<td>0</td>
<td>92</td>
<td>0</td>
<td>83.67</td>
<td>0.58</td>
<td>58.33</td>
<td>0.58</td>
<td>28.67</td>
<td>0.58</td>
<td>28.67</td>
</tr>
<tr>
<td>455</td>
<td>5 0</td>
<td>22</td>
<td>0</td>
<td>47.33</td>
<td>0.58</td>
<td>71.33</td>
<td>0.58</td>
<td>94.67</td>
<td>0.58</td>
<td>77.33</td>
<td>0.58</td>
<td>55</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>405</td>
<td>7 0</td>
<td>24.33</td>
<td>0.58</td>
<td>47</td>
<td>0</td>
<td>71</td>
<td>0</td>
<td>94.67</td>
<td>0.58</td>
<td>77.67</td>
<td>0.58</td>
<td>54.33</td>
<td>0.58</td>
<td>30.33</td>
<td>1.15</td>
<td>30.33</td>
</tr>
<tr>
<td>355</td>
<td>4 0</td>
<td>22</td>
<td>0</td>
<td>47.33</td>
<td>0.58</td>
<td>73</td>
<td>0</td>
<td>94.67</td>
<td>0.58</td>
<td>78.67</td>
<td>0.58</td>
<td>54</td>
<td>0</td>
<td>29.67</td>
<td>0.58</td>
<td>29.67</td>
</tr>
<tr>
<td>305</td>
<td>5 0</td>
<td>22</td>
<td>0</td>
<td>45.67</td>
<td>0.58</td>
<td>74</td>
<td>0</td>
<td>93</td>
<td>0</td>
<td>79</td>
<td>0</td>
<td>51.67</td>
<td>0.58</td>
<td>26.67</td>
<td>0.58</td>
<td>26.67</td>
</tr>
<tr>
<td>280</td>
<td>7 0</td>
<td>17.67</td>
<td>0.58</td>
<td>45.67</td>
<td>0.58</td>
<td>75</td>
<td>0</td>
<td>92</td>
<td>0</td>
<td>81</td>
<td>0</td>
<td>53.33</td>
<td>0.58</td>
<td>26</td>
<td>0</td>
<td>26</td>
</tr>
<tr>
<td>255</td>
<td>4 0</td>
<td>20.67</td>
<td>0.58</td>
<td>46</td>
<td>0</td>
<td>72.33</td>
<td>0.58</td>
<td>96</td>
<td>0</td>
<td>76.67</td>
<td>0.58</td>
<td>54</td>
<td>0</td>
<td>28.67</td>
<td>0.58</td>
<td>28.67</td>
</tr>
<tr>
<td>230</td>
<td>3 0</td>
<td>22</td>
<td>0</td>
<td>45</td>
<td>0</td>
<td>71.33</td>
<td>0.58</td>
<td>95.33</td>
<td>0.58</td>
<td>78</td>
<td>0</td>
<td>54</td>
<td>0</td>
<td>27</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>205</td>
<td>4 0</td>
<td>22</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>74</td>
<td>0</td>
<td>95</td>
<td>0</td>
<td>81</td>
<td>0</td>
<td>55</td>
<td>0</td>
<td>28</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>180</td>
<td>5 0</td>
<td>21.67</td>
<td>0.58</td>
<td>48</td>
<td>0</td>
<td>75</td>
<td>0</td>
<td>95</td>
<td>0</td>
<td>82</td>
<td>0</td>
<td>57</td>
<td>0</td>
<td>29</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>155</td>
<td>7 0</td>
<td>21</td>
<td>0</td>
<td>48</td>
<td>0</td>
<td>76</td>
<td>0</td>
<td>95</td>
<td>0</td>
<td>84</td>
<td>0</td>
<td>58</td>
<td>0</td>
<td>28</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>130</td>
<td>8 0</td>
<td>24</td>
<td>0</td>
<td>47.67</td>
<td>0.58</td>
<td>74</td>
<td>0</td>
<td>94</td>
<td>0</td>
<td>82</td>
<td>0</td>
<td>58.67</td>
<td>0.58</td>
<td>30.67</td>
<td>0.58</td>
<td>30.67</td>
</tr>
<tr>
<td>105</td>
<td>6 0</td>
<td>22</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>72</td>
<td>0</td>
<td>94.33</td>
<td>0.58</td>
<td>79</td>
<td>0</td>
<td>55.33</td>
<td>1.15</td>
<td>29</td>
<td>0</td>
<td>29</td>
</tr>
</tbody>
</table>

| Grand Mean/ sd | 6.67 | 2.23 | 23.19 | 4.96 | 46.58 | 0.92 | 72.32 | 3.21 | 92.81 | 2.53 | 78.51 | 5.02 | 55.25 | 1.78 | 29.72 | 3.12 |
Table 2.2 Nasalance scores for square wave test sounds in different stereo panoramas, recorded with the Nasometer 6450.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Left Mean</th>
<th>SD</th>
<th>3/4 Left Mean</th>
<th>SD</th>
<th>Balanced Mean</th>
<th>SD</th>
<th>3/4 Right Mean</th>
<th>SD</th>
<th>Right Mean</th>
<th>SD</th>
<th>Rev 3/4 Left Mean</th>
<th>SD</th>
<th>Rev Balanced Mean</th>
<th>SD</th>
<th>Rev 3/4 Right Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>805</td>
<td>9</td>
<td>0</td>
<td>24</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>70.33</td>
<td>0.58</td>
<td>94</td>
<td>0</td>
<td>77</td>
<td>0</td>
<td>52</td>
<td>0</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>755</td>
<td>7.33</td>
<td>0.58</td>
<td>26</td>
<td>0</td>
<td>47</td>
<td>0</td>
<td>70.33</td>
<td>0.58</td>
<td>93.67</td>
<td>0.58</td>
<td>72.33</td>
<td>0.58</td>
<td>51</td>
<td>0</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>705</td>
<td>6.33</td>
<td>0.58</td>
<td>23</td>
<td>0</td>
<td>47</td>
<td>0</td>
<td>72</td>
<td>0</td>
<td>93</td>
<td>0</td>
<td>76</td>
<td>0</td>
<td>52</td>
<td>0</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>655</td>
<td>6</td>
<td>0</td>
<td>22</td>
<td>0</td>
<td>44.67</td>
<td>0.58</td>
<td>71</td>
<td>0</td>
<td>95</td>
<td>0</td>
<td>77</td>
<td>0</td>
<td>53</td>
<td>0</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>605</td>
<td>9</td>
<td>0</td>
<td>21</td>
<td>0</td>
<td>45</td>
<td>0</td>
<td>72</td>
<td>0</td>
<td>93</td>
<td>0</td>
<td>79</td>
<td>0</td>
<td>53</td>
<td>0</td>
<td>26.67</td>
<td>0.58</td>
</tr>
<tr>
<td>555</td>
<td>7.33</td>
<td>0.58</td>
<td>20</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>74</td>
<td>0</td>
<td>92</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>52</td>
<td>0</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>505</td>
<td>6</td>
<td>0</td>
<td>17.33</td>
<td>0.58</td>
<td>42</td>
<td>0</td>
<td>69</td>
<td>0</td>
<td>96.67</td>
<td>0.58</td>
<td>77</td>
<td>0</td>
<td>54</td>
<td>0</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>455</td>
<td>4</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>71</td>
<td>0</td>
<td>95.33</td>
<td>0.58</td>
<td>79.33</td>
<td>0.58</td>
<td>55</td>
<td>0</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>405</td>
<td>10</td>
<td>0.58</td>
<td>22</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>73</td>
<td>0</td>
<td>91.33</td>
<td>0.58</td>
<td>76</td>
<td>0</td>
<td>53</td>
<td>0</td>
<td>28</td>
<td>0</td>
</tr>
<tr>
<td>355</td>
<td>5</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>45</td>
<td>0</td>
<td>75</td>
<td>0</td>
<td>92</td>
<td>0</td>
<td>81</td>
<td>0</td>
<td>53</td>
<td>0</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>305</td>
<td>7</td>
<td>0</td>
<td>18</td>
<td>0</td>
<td>42</td>
<td>0</td>
<td>70</td>
<td>0</td>
<td>95</td>
<td>0</td>
<td>76</td>
<td>0</td>
<td>50.67</td>
<td>0.58</td>
<td>24.67</td>
<td>0.58</td>
</tr>
<tr>
<td>255</td>
<td>8</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>73</td>
<td>0</td>
<td>96</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>53</td>
<td>0</td>
<td>24.33</td>
<td>0.58</td>
</tr>
<tr>
<td>205</td>
<td>8</td>
<td>0</td>
<td>19.33</td>
<td>0.58</td>
<td>46</td>
<td>0</td>
<td>74</td>
<td>0</td>
<td>93</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>54</td>
<td>0</td>
<td>26</td>
<td>0</td>
</tr>
<tr>
<td>155</td>
<td>7.33</td>
<td>0.58</td>
<td>18</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>75</td>
<td>0</td>
<td>93</td>
<td>0</td>
<td>81</td>
<td>0</td>
<td>55</td>
<td>0</td>
<td>26</td>
<td>0</td>
</tr>
<tr>
<td>130</td>
<td>7.33</td>
<td>0.58</td>
<td>17</td>
<td>0</td>
<td>45</td>
<td>0</td>
<td>74</td>
<td>0</td>
<td>91.33</td>
<td>0.58</td>
<td>81</td>
<td>0</td>
<td>55</td>
<td>0</td>
<td>26</td>
<td>0</td>
</tr>
<tr>
<td>105</td>
<td>8.33</td>
<td>0.58</td>
<td>19</td>
<td>0</td>
<td>44</td>
<td>0</td>
<td>71</td>
<td>0</td>
<td>94.33</td>
<td>0.58</td>
<td>78</td>
<td>0</td>
<td>53</td>
<td>0</td>
<td>27</td>
<td>0</td>
</tr>
</tbody>
</table>

Grand Mean/ sd  7.26  1.42  20.30  2.55  45.35  1.43  72.51  1.91  93.58  1.55  78.77  2.66  53.46  1.64  26.35  1.70
A repeated-measures ANOVA was run for loudspeaker panorama (eight loudspeaker stereo panoramas), nasometer model (6200 vs. 6450) and repetition (three repetitions). Where sphericity was violated, the Greenhouse-Geisser correction was used. There were significant effects for panorama \( (p < .001, F(1.465,126) = 3192.752) \), nasometer \( (p < .005, F(1,18) = 18.419) \) and the panorama-nasometer interaction \( (p < .01, F(2.159, 38.857) = 5.836) \). The effect of repetition and the remaining interactions were not significant.

Post hoc t-tests demonstrated that the mean of the nasalance scores from the Nasometer 6200, 50.629 (SD 28.083) was about one point greater than the mean of 49.697 (SD 28.735) for the Nasometer 6450 \( (p < .01) \). When averaged across Nasometer models, the mean nasalance scores for each of the loudspeaker panoramas, including the three with the microphones reversed, were significantly different from each other \( (all \ p < .01) \).

With regards to the panorama-Nasometer interaction effect, the amount by which the Nasometer 6200 scored higher than the Nasometer 6450 depended on the loudspeaker panorama for the square wave playback. The nasometer differences between the stereo panoramas for three-quarter left, balanced, balanced (reversed) and three-quarter right (reversed) were significant \( (p < .01) \).

### 2.3 Study 2: Comparative Assessment of the Nasometers 6200 and 6450 with Normal Participants

#### 2.3.1 Participants

Twenty-five females were recruited from the student population of the University of Toronto’s Speech-Language Pathology program. The study was limited to female participants because they were readily available, comprising the majority of the program’s student body, as is typical for a
speech language pathology program (Boyd & Hewlett, 2001). Whitehill’s (2001) study was also conducted with all female subjects. The participants were between 22 and 30 years of age (M = 24.16, SD = 2.375) and spoke English with the accent that is common to Southern Ontario. Each participant completed six recording sessions between May and August 2011. Twenty-one of the participants completed their sessions within 7 weeks from start of the study. Due to scheduling conflicts, four participants’ recordings took up to 11 weeks to complete. The average time between the first and the final recording was 35 days (range 16 to 77 days).

2.3.2 Methods

All the nasalance measurements took place in the same sound-treated therapy room. The Nasometer 6200 was connected to a computer running the Windows 95 operating system. The Nasometer II 6450 was connected to a computer running the Windows XP operating system. The Nasometers were calibrated according to the manufacturers’ specifications, at the beginning of each day. In each session, the participants were assigned to begin with either the Nasometer 6200 or the Nasometer 6450, according to a previously determined randomization schedule. The order of the stimuli was also randomized.

We used oral, balanced (oral-nasal) and nasal stimuli for the recordings of the participants on the two Nasometers. The stimuli were abbreviated versions of the Zoo Passage (Fletcher, 1976), the Rainbow Passage (Fairbanks, 1963) and the Nasal Sentences (Fletcher, 1976). We used the first two sentences of the Zoo Passage (“Look at this book with us. It’s a story about a zoo.”), the second sentence of the Rainbow Passage (“The rainbow is a division of white light into many beautiful colours.”) and the first sentence from the Nasal Sentences (“Mama made some lemon jam.”). In a previous comparison, no difference was found between the full and abbreviated
nasal stimuli, but the scores for the abbreviated oral stimulus were two points lower than the full Zoo Passage (Bressmann, 2005). The abbreviated versions were used to obtain as many measurements as possible in the limited time the student participants were available.

### 2.3.3 Statistical Analysis

The nasalance scores of the Nasometers were compared with a repeated measures ANOVA and post-hoc t-tests. The test-retest variability between measurements was examined by calculating the absolute differences in scores between sessions and creating cumulative frequency tables. As the absolute differences in nasalance scores were not normally distributed, the Wilcoxon signed-rank test was used to determine if some stimuli had greater variance than others. To assess the relationship between the subjects’ nasalance scores and their mean test-retest variability, we calculated Pearson product-moment correlation coefficients.

### 2.3.4 Results

The mean nasalance scores were calculated by averaging the scores for each stimulus for each Nasometer over the six sessions. The results are displayed in Table 2.3.

A repeated-measures ANOVA of the nasalance scores was run for the three stimuli, by two nasometers over six sessions. The ANOVA showed an effect for nasometer model F(1,24) = 4.419, p < .05; for session F(5,120) = 2.629, p < .05; for stimulus F(1.426, 34.235) = 2285.009, p < .01; and a nasometer-stimulus interaction effect F(2,48) = 15.307, p < .01. There were no interaction effects for nasometer-session, session-stimulus or nasometer-session-sentence. As
Maulchy’s test of sphericity was positive for the stimulus effect, the Greenhouse-Geisser correction was applied. The nature of the main effects was further explored with post hoc paired t-tests.

Paired t-tests were used to compare the two nasometers post hoc. The mean nasalance score for the Nasometer 6450 (36.47, SD 21.686) was significantly higher (p < .05) than the mean of the Nasometer 6200 (35.56, SD 20.885). The mean nasalance scores of the oral, balanced and nasal stimuli (13.16, SD 6.121; 32.22, SD 5.800; and 62.67, SD 6.267, respectively) differed significantly (all differences p < .01). For the sentence-nasometer interaction effect, paired t-tests showed that the Nasometer 6450 scored significantly higher than the Nasometer 6200 for the balanced stimulus (p < .05) and the nasal stimulus (p < .01). For the effect of session, the paired t-tests indicated that the mean nasalance score of the sixth session (37.10, SD 5.090) differed significantly from the first (35.43, SD 5.308), second (35.560, SD 5.364) and third sessions (35.66, SD 4.305) (all differences p < .05).
### Table 2.3 Mean nasalance scores and standard deviation by Nasometer model and session.

<table>
<thead>
<tr>
<th>Session</th>
<th>Oral Mean</th>
<th>Oral SD</th>
<th>Balanced Mean</th>
<th>Balanced SD</th>
<th>Nasal Mean</th>
<th>Nasal SD</th>
<th>Oral Mean</th>
<th>Oral SD</th>
<th>Balanced Mean</th>
<th>Balanced SD</th>
<th>Nasal Mean</th>
<th>Nasal SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.72</td>
<td>4.64</td>
<td>31.28</td>
<td>5.14</td>
<td>61.28</td>
<td>6.66</td>
<td>12.28</td>
<td>7.01</td>
<td>32.88</td>
<td>6.88</td>
<td>62.92</td>
<td>6.79</td>
</tr>
<tr>
<td>2</td>
<td>12.6</td>
<td>5.22</td>
<td>31.12</td>
<td>4.7</td>
<td>61.32</td>
<td>5.26</td>
<td>12.6</td>
<td>5.35</td>
<td>32.8</td>
<td>5.66</td>
<td>63.52</td>
<td>5.81</td>
</tr>
<tr>
<td>3</td>
<td>13.12</td>
<td>6.77</td>
<td>31.48</td>
<td>5.78</td>
<td>61.76</td>
<td>6.37</td>
<td>12.84</td>
<td>6.18</td>
<td>32.36</td>
<td>5.82</td>
<td>63.84</td>
<td>5.67</td>
</tr>
<tr>
<td>4</td>
<td>14.12</td>
<td>7.16</td>
<td>32.08</td>
<td>6.77</td>
<td>62.52</td>
<td>6.6</td>
<td>13.76</td>
<td>6.35</td>
<td>32.72</td>
<td>5.79</td>
<td>63.56</td>
<td>6.44</td>
</tr>
<tr>
<td>5</td>
<td>14.68</td>
<td>6.37</td>
<td>32.88</td>
<td>5.1</td>
<td>62.36</td>
<td>6.22</td>
<td>14.4</td>
<td>6.36</td>
<td>33.84</td>
<td>5.96</td>
<td>64.44</td>
<td>6.45</td>
</tr>
</tbody>
</table>

Grand Mean / SD

<table>
<thead>
<tr>
<th>Oral Mean</th>
<th>Oral SD</th>
<th>Balanced Mean</th>
<th>Balanced SD</th>
<th>Nasal Mean</th>
<th>Nasal SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.21</td>
<td>5.904</td>
<td>31.68</td>
<td>5.468</td>
<td>61.81</td>
<td>6.164</td>
</tr>
<tr>
<td>13.12</td>
<td>6.349</td>
<td>32.75</td>
<td>6.085</td>
<td>63.54</td>
<td>6.269</td>
</tr>
</tbody>
</table>
The variability of nasalance scores over time was assessed by calculating the differences in scores between the six sessions for each sentence, nasometer and participant. The differences were converted to absolute values, and the mean and standard deviation for each Nasometer-sentence combination were calculated. These results, along with the cumulative frequencies, are displayed in Table 2.4. Wilcoxon signed-rank tests were conducted to compare the mean differences for the different stimuli. There were significant differences between the nasal stimulus and the oral \( z = -2.678, p < .01 \) and balanced stimuli \( z = -2.886, p < .01 \) for the Nasometer 6200. For the Nasometer 6450, there was a significant difference between the nasal and oral stimuli \( z = -3.116, p < .01 \). There were no significant differences between the machines.

The cumulative frequencies of the absolute differences in nasalance points across six sessions are also shown in Table 2.4. For both Nasometers, 90.7% of the participants’ nasalance scores for the oral stimulus were within 6 points. For the balanced sentence, 92.3% (Nasometer 6200) and 93.3% (Nasometer 6450) of the scores were within 7 points of the participants’ other scores. An 8 point spread captured 91.5% (Nasometer 6200) and 91.7% (Nasometer 6450) of nasalance scores for the nasal sentence.

To assess whether participants with higher nasalance scores had higher test-retest variability, the mean values of the participants’ nasalance scores were correlated with the mean of their differences between sessions. Significant correlations were found for the oral stimulus for the Nasometer 6200 \( r = .733, p < .01 \) and for the Nasometer 6450 \( r = .543, p < .01 \).
Table 2.4. Means and cumulative frequencies of test-retest differences by Nasometer model

<table>
<thead>
<tr>
<th>Nasalance Score Difference</th>
<th>Nasometer 6200</th>
<th>Nasometer 6450</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Oral Balanced Nasal</td>
<td>Oral Balanced Nasal</td>
</tr>
<tr>
<td>0-5</td>
<td>315 / 84.0</td>
<td>309 / 82.4</td>
</tr>
<tr>
<td>≤ 6</td>
<td><strong>340 / 90.7</strong></td>
<td>332 / 88.5</td>
</tr>
<tr>
<td>≤ 7</td>
<td>347 / 92.5</td>
<td><strong>346 / 92.3</strong></td>
</tr>
<tr>
<td>≤ 8</td>
<td>354 / 94.4</td>
<td>359 / 95.7</td>
</tr>
<tr>
<td>≤ 9</td>
<td>357 / 95.2</td>
<td>370 / 98.7</td>
</tr>
<tr>
<td>≤ 14</td>
<td>375 / 100</td>
<td>375 / 100</td>
</tr>
<tr>
<td>≤ 19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean difference

<table>
<thead>
<tr>
<th></th>
<th>Nasometer 6200</th>
<th>Nasometer 6450</th>
</tr>
</thead>
<tbody>
<tr>
<td>difference</td>
<td>2.89</td>
<td>2.99</td>
</tr>
<tr>
<td>SD</td>
<td>1.51</td>
<td>1.163</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Nasometer 6200</th>
<th>Nasometer 6450</th>
</tr>
</thead>
<tbody>
<tr>
<td>difference</td>
<td>2.83</td>
<td>3.20</td>
</tr>
<tr>
<td>SD</td>
<td>1.191</td>
<td>1.122</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Nasometer 6200</th>
<th>Nasometer 6450</th>
</tr>
</thead>
<tbody>
<tr>
<td>difference</td>
<td>3.92</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.109</td>
<td></td>
</tr>
</tbody>
</table>
2.4 Discussion

The current study had the goal of characterizing the new Nasometer 6450 in comparison to the older model 6200 using both synthetic test sounds and control participants. A particular focus of the investigation was on the test-retest variability of the two instruments.

The initial analysis of the sound pressure levels transmitted by the two loudspeakers for the synthetic test sounds demonstrated that the left loudspeaker was slightly louder. These differences were probably due to minor differences in the manufacturing of the two loudspeakers. For the main effect of stereo panorama (averaged across the two machines), significant differences between all panoramas were expected and confirmed. However, three panoramas (¾ right, balanced and ¾ left) were also tested with the loudspeaker-microphone alignment reversed. Here, the significant differences between the mean for the panorama and its respective reversed counterpart were not expected. The increased sound pressure from the left loudspeaker is the likely cause of these significant differences.

The mean nasalance scores obtained from the hard left and hard right sound files, 7 and 93%, respectively, suggest that 7% of the acoustic energy presented on one side of the sound separator plate makes its way to the opposite microphone. According to the manufacturer, the sound separator plate attenuates sound transmission to the opposite microphone by 25 dB (KayPENTAX, 2010). Therefore, the nasalance scores will never be zero or 100 (Gildersleeve-Neumann & Dalston, 2001). In addition, the mean values for the loudspeaker balances only roughly approximate the theoretically expected scores of 0%, 25%, 50%, 75% and 100% nasalance for the hard left (oral microphone) three-quarter left, balanced, three-quarter right, and hard right (nasal microphone) stereo panoramas, respectively.
With the test frequency of 755 Hz, our Nasometer 6200 produced nasalance scores that deviated from the scores for the other frequencies. This effect was reproducible and persisted when the microphones were reversed. Since the same effect was not observed for the Nasometer 6450, it did not appear that this particular frequency excited resonances in any elements of the recording contraption. It would be interesting to study whether this effect was specific to our particular Nasometer 6200 or if other instruments show a similar phenomenon for this frequency. Zajac et al. (1996) demonstrated differences in the measurements obtained with different headsets for the Nasometer 6200. Both Nasometers were used with the headsets that they were delivered with, so the possibility of a mismatch in microphone sensitivity cannot be ruled out.

The analysis of the data for the normal participants demonstrated that means of the nasalance scores for the oral, balanced and nasal stimuli for the Nasometer 6200 were between 2 to 5 points greater than those previously reported by Bressmann (2005). We assessed the differences between this study and the present study with pooled t-tests and found a significant difference for the nasal stimulus (p < .01). Since the Nasometer 6200 used was the same in both studies, we conclude that the participants must have been the source of variability.

While the differences in mean nasalance scores obtained by the Nasometers 6200 and 6450 were significant, it is unlikely that their magnitude would affect clinical practice. The results of the research overall confirm the manufacturer’s claim that the Nasometers 6200 and 6450 score within two points of each other. The ANOVA’s significant sentence*machine interaction effect demonstrated that the difference in scores between the machines was dependent on the sentence. The post-hoc tests showed that the Nasometer 6450 scored higher as the proportion of nasal
consonants increased. This two-point difference, albeit small, should be taken into consideration when comparing the nasalance scores obtained from test materials loaded with nasal consonants.

The significant effect of session was not anticipated. The recording location and the procedures were kept constant. Participants were informally screened for allergies or upper respiratory tract infections before each recording. Only one participant developed a cold during the study, so she was asked to hold off on her recordings until the infection had cleared up. While one could speculate that familiarity with the material may have led to faster speaking rates over time, Gauster, Yunusova, & Zajac (2010) demonstrated that speaking rate does not affect nasalance scores. We suspect that the environmental temperature may have had a bearing on the nasalance scores over time because it is known that air temperature can affect nasal patency (Olsson & Bende, 1985). The first few weeks of the study were rainy and cooler while the weather towards the end of the study was sunny and warm. However, we did not record the temperature or humidity of the room during the recording sessions.

The mean differences in scores between sessions for the oral and balanced stimuli were comparable to those found by Lewis et al. (2008) using the Nasometer 6400 while the mean differences for the nasal materials were comparable to Whitehill (2001). The trend of increasing variability of scores with increased nasal content that was evident in the data reported by Whitehill (2001) was confirmed in the present study. This increase may be a simple effect of the averaging of the nasalance score because a stimulus containing nasal sounds will have nasalance spikes in the nasalance trace. Such statistical outliers can skew the frequency distribution and distort the arithmetic mean (Ferguson, 1981). While there were significant differences between the variability of the stimuli, there were none between the nasometers.
For clinical practice, the distribution of the differences in nasalance points is more meaningful than the mean differences and standard deviations. The cumulative frequency tables illustrated an increase in measurement variability as the nasal content of the test stimuli increased. For both the Nasometer 6200 and the Nasometer 6450, a 6 point spread captured 90.7% of the differences in scores for the oral stimuli. A 7 point spread captured 92.3% of the differences for the balanced stimulus for the Nasometer 6200 and 93.3% for the Nasometer 6450. These results are very similar to what Lewis et al. (2008) found for these stimuli with the Nasometer 6400. For the nasal sentence, an 8 point spread accounted for 91.5% of the differences for the Nasometer 6200 and for 91.7% for the Nasometer 6450. This was comparable to the results by Whitehill (2001) who found that 89.3% of her participants were within 8 points. Given these results, clinicians would be advised to allow as much as 6 nasalance points for normal variation with an oral stimulus and 8 nasalance points for a nasal stimulus when using a nasometer to assess possible changes in a patient’s oral-nasal balance.

There were significant moderate to strong correlations when participants’ mean nasalance scores for the oral stimulus were correlated with their mean differences of scores between sessions. This may explain why the test-retest variability of individuals with cleft palate (Watterson et al., 2006) has been found to be greater than for normal participants (Lewis et al., 2008).

2.5 Conclusion

On average, nasalance scores measured with the new Nasometer model 6450 fell within two points of the Nasometer 6200, so the two instruments should yield comparable results in clinical practice. For both models, the variability in scores increased with the proportion of nasal consonants in the sentence. Depending on the phonetic content of the test materials, clinicians
should allow for 6-8 point between-session variability when interpreting nasalance scores. Participants with higher nasalance scores for oral stimuli tend to have higher between-session variability. The challenge of achieving a satisfactory level of diagnostic accuracy despite the inherent variability of nasalance scores should be addressed in future research.
3. Towards an Assessment of Resonance Disorders Based on Linear Discriminant Analysis

3.1 Introduction

Computerized nasometry, with instruments such as the Nasometer by KayPentax, is a common instrumental adjunct to the perceptual clinical evaluation of resonance disorders. While nasometry has also been used in the assessment of hearing impaired and dysarthric individuals (Whitehall & Lee, 2008), its primary role has been the assessment of hypernasality in patients with cleft lip and palate (Kummer, 2008). The nasalance score reflects the proportion of oral to nasal sound energy in speech and is calculated as follows: nasalance = nasal/(nasal+oral) x 100 (Fletcher, 1976). When there is excess nasal resonance, the scores for speech stimuli without nasal sounds are higher than normal, suggesting hypernasality. When there is a lack of nasal resonance, the scores are lower for speech stimuli loaded with nasal consonants, suggesting hyponasality (Kummer, 2008; Dalston et al., 1991a, 1991b).

Hypernasality is usually due to velopharyngeal dysfunction (structural or neurogenic), oronasal fistulae or mislearning, while hyponasality almost always result from an obstruction in the nasopharynx or the nasal cavities (Kummer, 2011). Mixed nasality comprises both hypernasality and hyponasality and can occur when there is both velopharyngeal dysfunction and a blockage such as a deviated septum from unilateral cleft lip and palate (Kummer 2008, Peterson-Falzone et al, 2001). Cul-de-sac resonance is an extreme form of mixed or de-nasality. It is also due to a blockage, but the sound is trapped in the nasal cavities and the speech sounds muffled (Kummer, 2011). The relationship between a resonance disorder and a nasalance score is less straightforward in patients who have mixed or cul-de-sac nasality. The nasalance scores are
expected to be normal or close to normal (Kummer, Billmire, & Myer, 1993) but there are only a few published reports with the nasalance scores of only four individuals published in the literature (Kummer et al., 1993, Karnell et al., 2004, Van Lierde, Luyten, Mortier, Tijskens, Bettens & Vermeersch, 2011). Peterson-Falzone et al. (2001) state that the instrumental assessment of mixed nasality is “long overdue” (p163).

The clinician’s trained ear is considered the gold standard in assessing resonance disorders (Kuehn & Moller, 2000; Moon, 2009; Whitehill & Lee, 2008). Yet, listener’s perceptual ratings have been described as difficult, subjective and of poor reliability (Keuning, Wieneke, & Dejonckere, 2004; Whitehill & Lee, 2008). Therefore, quantitative and objective corroboration from instrumentation has been sought (Kuehn & Moller, 2000). The Nasometer by Kay Pentax is the most widely used of the acoustic instruments. The various factors that affect nasalance scores and their correspondence with listeners’ perception, have been studied for decades.

In the clinical setting and most research studies, hypernasality and/or hyponasality are rated perceptually. Traditionally, ordinal scales with equal appearing intervals are used for the ratings. However, intra-rater and inter-rater variability is notoriously variable. A recent listener study of hypernasality and nasometry using a five point scale reported exact intra-listener agreement of 25-100% and exact inter-listener agreement between 33-62%. When agreement was broadened to be within one of the five points, the intra-listener agreement rose to 87.5-100%, and the inter-listener agreement rose to 67-96%. The higher agreements were reached between speech–language pathologists working in cleft palate centres (Brunnegård et al., 2012). A study employing a visual analog scale with an anchor sample reported a Spearman correlation for hypernasality of $r_s = .49$, although this was slightly higher ($r_s = .55$) among more experienced
speech-language pathologists (Keuning, Wieneke, Van Wijngaarden, & Dejonckere, 2002). Some studies have shown that experience (Brunnegård et al., 2012) or practice (Lee et al., 2009) leads to better agreement between raters, while others have not (Keuning et al., 2002, Lewis, Watterson, & Houghton, 2003). Part of the problem may be the scales themselves. Whitehill, Lee & Chun (2002) suggested that hypernasality is a prothetic continuum, whereby “[…] listeners have difficulty partitioning hypernasality into equal-appearing intervals because of the psychophysical nature of the dimension” (p.85). With direct magnitude estimation, a value is assigned to an initial stimulus and the listeners assign subsequent stimuli numbers in proportion to the first. The authors found perceptual judgments based on direct magnitude estimation more valid and reliable (Whitehill et al., 2002).

A nasalance score is meant to aid in distinguishing normal from disordered resonance. Ideally, it should also allow the clinician to gauge severity. However, many factors influence oral nasal resonance and the resulting nasalance score. A non-exhaustive list includes: the sounds spoken, vowel content (Awan et al., 2011, Lewis, Watterson & Quint, 2000), language (Dalston, Neiman & Gonzalez-Landa, 1993; Van Lierde, Wuyts, De Bodt, & Van Cauwenberge, 2001, Whitehill, 2001), regional accents (Dalston et al., 1993; Seaver et al., 1991), pitch (Van Lierde et al., 2011), nasal congestion (Birkent, Erol, Ciyiltepe, Eadie, Durmaz, & Tosun, 2009; Pegoraro-Krook, Dutka-Souza, Williams, Teles Magalhaes, Rossetto, & Riski, 2006; Williams, Eccles, & Hutchings, 1990), audible nasal emissions (Dalston et al., 1991b; Karnell, 1995), instrument (Awan et al., 2011, Bressmann, 2005; Bressmann et al., 2006; Lewis & Watterson, 2003; Watterson et al., 2005), and simple random/daily variation (Chapter 2; Lewis et al., 2008; Watterson & Lewis, 2006; Whitehill 2001). For the researchers and clinicians around the world interested in what scores to expect locally with their instrument, multiple normative values of
standardized passages have been published to reflect regional and instrumental differences (Anderson, 1996; Awan et al., 2011; Bressmann, 2005; Bressmann et al., 2006; Brunnegård & van Doorn, 2009; Dalston et al., 1993; Haapanen, 1991; Hogen Esch & Dejonckere, 2004; Kavanagh, Fee, & Kalinowski, 1994; Lewis & Watterson, 2003; Lewis et al., 2000; Nichols, 1999; Tachimura, Mori, Hirata, & Wada, 2000; Seaver et al., 1991; Van Doorn & Purcell, 1998; Van Lierde et al., 2001; Watterson et al., 2005; Whitehill, 2001).

The relationship between nasality ratings and nasalance scores is commonly evaluated in two ways. Sensitivity and specificity are used to find an appropriate cut off score that best distinguishes normal from disordered resonance. Correlation analyses are used to describe the relationship between nasalance scores and perceptual evaluations of nasality (be they small range scales or direct magnitude estimations). Hypernasality is commonly assessed with a text passage without nasal speech sounds, such as the Zoo Passage (Fletcher, 1976). The best results for sensitivity and specificity were found with cut off scores between 26 and 32 with overall efficiencies between .69 and .87 (Dalston et al., 1991a, 1993; Hardin et al., 1992). Hyponasality is commonly assessed using a text passage loaded with nasal consonants. Nasalance scores below 50 have reported sensitivities (in detecting presence or absence of hyponasality) of .48 to 1.00 and reported specificities (to exclude normal resonance) between .79 and .91 (Dalston et al., 1991b; Hardin et al., 1992). Dalston et al. (1991) point out that the inclusion or exclusion of participants with pharyngeal flaps or audible nasal emissions may influence the range of reported sensitivities and specificities.

In some studies, the correlation between the perception of the severity of a hypernasal resonance disorder and nasalance scores was found to be moderate to strong, which values such as $r_s = .62$ to
Likewise, the perception of hyponasality and nasalance scores also correlated moderately to strongly, \( r_s = -.65 \) (Karnell et al., 2004) and \( r = -.76 \) (Sweeney & Sell, 2008). Lower or insignificant correlations between hypernasality and nasalance scores have been found for inexperienced listeners (Nellis, Neiman & Lehman, 1992; Brunnegård et al., 2012), as well as for experienced speech language pathologists, \( r_s = .36 \) to \( .60 \) (Keuning et al., 2002). For this last study, it should be noted that the agreement between the listeners was not very strong to begin with (\( r = .49 \)). The agreement between the listeners was only slightly better than between the listeners and the nasalance scores. The use of direct magnitude estimation of hypernasality instead of equal appearing interval scales does not seem to improve the agreement between listener perceptions of nasality and nasalance scores (Brancamp et al., 2010).

Some efforts have been made to use the nasalance scores differently. The nasalance distance was introduced by Bressmann et al. (2000, 2006). The measure is the difference between the score from a nasally loaded stimulus and that of an oral stimulus. It reflects how much distinction the speaker can make between oral and nasal sounds. The nasalance distance was shown to have greater sensitivity and specificity for the perception of hypernasality than the magnitude of the oral stimulus alone (Bressmann et al., 2006). Nasalance scores have also been used in conjunction with other measures. The Nasality Severity Index was introduced by Van Lierde, Wuyts, Bonte, & Van Cauwenberge (2007) to aide in the diagnosis of mild hypernasality. It includes a perceptual evaluation along with aerodynamic measurements, nasometry and the Glatzel test (which measures condensation on a mirror placed below the nostril) in order to create an objective measure that “reflects the multidimensional nature of resonance” (Van Lierde et al., 2007). However, this approach has not found much use in research or clinical practice since.
The overwhelming majority of nasometry studies have focused on hypernasality. Hypernasality affects speech intelligibility and acceptability more than hyponasality (Shprintzen, Lewin, & Croft., 1979), so it is clinically more relevant. However, the attempt to fit every patient diagnostically along a one-dimensional continuum of hypernasality may overlook the possible range of individual variability. Hypernasal resonance in combination with compromised nasal patency is common among patients with bilateral and unilateral complete cleft lip and palate (Fukushiro & Trindale, 2005) due to combinations of a deviated septum, a narrow vestibule and hypertrophic turbinates (Coston, Friedman & VanLue, 2009). Some diagnostic schemes, such as the Great Ormond Street Assessment '98 (Sell, Harding & Grunwell, 1999) include rating scales for both hyper- and hyponasality. This allows the clinician to document separate observations and comment on both aspects of the resonance disorder. However, with the exception of one case study, nobody has explicitly examined the nasalance scores associated with mixed resonance. In the case study of a patient with hypertrophic tonsils, Kummer et al. (1993) suggested the scores of cul-de-sac and mixed resonance were close to normal as the effects of hypernasality and hyponasality would cancel each other out.

In the present study, we argue that at least a part of the disagreement between listeners, or listeners and nasalance scores, might be attributed to mixed nasality. Clinically, mixed nasality may present as mild hypernasality because the hypernasality is attenuated by the hyponasality. As an example, we reviewed the data from Bressmann et al.’s nasalance distance study (2006). Table 3.1 shows that the mild hypernasality group had a mean nasalance score of 40.94 for the nasal stimulus, well below the cut off of 50 for hyponasality and well below the normal controls and the moderately hypernasal patients. This demonstrates that when the focus is simply hypernasality, the picture will often be incomplete. In the current study, we argue that a proper
diagnostic procedure should evaluate the nasalance scores for oral and nasal stimuli together to arrive at a unique classification for the speaker. Such a classification would be the first step towards an improved assessment of resonance disorders, based on nasometry scores.

**Table 3.1 Mean nasalance scores and nasalance distance with standard deviations for the Zoo Passage and the Nasal Sentences as measured with the Nasometer 6200 from Bressmann et al. (2006)**

<table>
<thead>
<tr>
<th></th>
<th>Zoo Passage (oral)</th>
<th>Nasal Sentences</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Normal speakers</td>
<td>50</td>
<td>13.45</td>
<td>5.94</td>
</tr>
<tr>
<td>Mild hypernasality</td>
<td>8</td>
<td>17.68</td>
<td>9.59</td>
</tr>
<tr>
<td>Moderate hypernasality</td>
<td>11</td>
<td>34.06</td>
<td>18.96</td>
</tr>
</tbody>
</table>

The goal of this study was to assess the influence of nasal patency, that is, how open the nasal passages are, on nasalance scores, using normal speakers simulating hyponasality, hypernasality and mixed resonance. Rather than providing diagnostic cutoff values, we used linear discriminant analysis to derive a formula that combined nasalance values for both oral and nasal stimuli to assign a speaker to a pre-determined diagnostic category.

The first hypothesis was that the hyponasal condition would have lower scores than the normal condition (and that the decrease would be greater for the nasal stimulus than the oral stimulus).
The second hypothesis was that the hypernasal condition would have higher scores than the normal condition (and that the increase would be greater for the oral stimulus than the nasal stimulus). The third hypothesis was that the mixed condition (hypernasal resonance with one nostril occluded) would yield normal scores as, per Kummer et al.’s (1993) prediction, the effects of hypernasality and hyponasality would cancel each other out. The first null hypothesis to be refuted was that the mean nasalance scores for simulated resonance disorders would not differ significantly from the normal speaking condition. Refuting this hypothesis would corroborate the success of the hypernasal and hyponasal simulations. Also, if the speaking conditions produced scores that were significantly different from each other, for at least one of the two stimuli, the nasalance scores could then be used in a linear discriminant analysis.

The fourth hypothesis was that the Nasalance distance for the normal condition would be greater than that of the hypernasal, hyponasal or mixed condition. The null hypothesis to be refuted was that the mean nasalance distance would not differ between conditions. It was expected from Bressmann et al.’s study (2006) that the hypernasal condition would have a smaller nasalance distance than the normal condition. Based on a study by Gildersleeve-Neumann & Dalston (2001), blocking both nostrils dropped the scores on the oral stimulus to the minimum, obstructing one nostril should decrease the scores of the oral or the nasal stimuli to some extent. There were no predictions regarding mixed resonance.
3.2 Methods

3.2.1 Participants

The recording sessions took place between September and October 2012. Sixteen normal speaking females were recruited from the student population at the University of Toronto. The participants were between 22 and 30 years of age (mean 24.1, SD 2.2) and spoke English with the accent that is common to Southern Ontario. They reported normal hearing, no history of cleft lip/palate, resonance disorder or excess nasal congestion.

3.2.2 Participant Training

The participants’ experimental task was to simulate different resonance disorders. As a first step, the first author explained the nature of different resonance disorders and practiced with the participants how to produce the associated voice qualities. The following types of oral-nasal balance were discussed and practised with the participants:

- Normal voice. This voice quality was discussed with the participants but no practice was necessary.
- Hypernasality was simulated by lowering the velum and nasalizing all speech sounds.
- Hyponasality was simulated by closing one nostril with the index finger. It is estimated that up to 80% of individuals experience a nasal cycle, whereby one nostril is more patent than the other at various times throughout the day (Hixon, Hoit & Weismer, 2008; Principato & Osenberger, 1970; Stoksted, 1953). To compensate for this the hyponasal condition was repeated for both nostrils so that the higher and lower patency nostrils could be identified.
Mixed nasality was simulated by speaking with a lowered the velum and one closed nostril. Like the hyponasal condition, this speaking condition was repeated to identify the higher and lower patency nostrils.

The author taught the participants hypernasal resonance by first explaining resonance disorders and then demonstrating them. The participants were asked to produce various nasal and non-nasal sounds, sustain nasalized vowels and repeat words and sentences with a voluntarily lowered velum. The participants were then given time to practice their hypernasal resonance with the test stimuli before the recordings. Additional time was given to practice and further demonstration was provided as required. Hyponasality was achieved by placing an index finger firmly over one ala of the nose and closing the corresponding nostril. While this manoeuvre required no practice per se, the participants were given specific instructions and practise time for the mixed nasality.

### 3.2.3 Stimuli

The stimuli consisted of an oral and a nasal stimulus. The first two sentences of the Zoo Passage ("Look at this book with us. It’s a story about a zoo") and the first of the Nasal Sentences ("Mama made some lemon jam") were used (Fletcher, 1976). In previous research, Bressmann (2005) compared these abbreviated versions of the stimuli to the full versions. With the Nasometer 6200 he found that the mean for the Zoo sentence was about two points lower than the full version but no significant difference between the first of the Nasal Sentences and the full set (Bressmann, 2005). The abbreviated versions were used to obtain as many measurements as possible in the limited time the student participants were available. The order of the stimuli was randomized and they were read twice for each condition.
3.2.4 Recording Procedures

All the nasalance measurements and recordings took place in a quiet room. The Nasometer II 6450 was connected to a laptop computer running the Windows 7 operating system. The Nasometer was calibrated according to the manufacturers’ specifications, prior to each day’s data collection. The audio files and statistics were saved using the Nasometer software for each speaking condition.

Since the sound quality of the Nasometer is affected by a filtering algorithm centred at 500Hz with a 300Hz bandwidth, additional high quality audio recordings were made using a Q3 Handy Video Recorder (Zoom, Tokyo, Japan). The recordings were made with the device’s internal directional stereo microphone with a signal resolution of 16 bit and a sampling rate of 44.1 kHz. The recordings were saved as *.wav files.

3.2.5 Simulation Verification

To verify the accuracy of the participants’ portrayal of hypernasal resonance, both authors listened to the audio recordings of the sessions. A consensus decision was made as to whether or not the participants had successfully simulated hypernasal resonance. As a result of this qualitative verification step, five participants were excluded from the data analysis, leaving a total of 11 data sets in the study.

3.2.6 Data Analysis

The nasalance values were analysed using SPSS 20.0. Descriptive measures were used to display the nasalance scores for the different speaking conditions. The impact of nasal patency on
nasalance scores were assessed with a repeated measures ANOVA. Due to nasal cycling, the nasal patency may be uneven between the two nostrils (Hixon et al., 2008). For the data analysis, the nasalance scores from blocking of the right or left nostril (hyponasal and mixed) were individually recoded into a higher and a lower patency nostril (less and more blocked respectively) for every speaker based on the magnitude of the two nasalance scores for each nostril. This was done to capture the effects of lesser and greater degrees of blockage which would have been lost when averaged across right and left.

Bressmann et al. (2000, 2006) suggest a measure of nasalance distance that is calculated by subtracting the nasalance score for an oral stimulus from the nasalance score a nasal stimulus. Nasalance distances were calculated for all speaking conditions and their magnitudes were compared using a one-way ANOVA with paired t-tests post hoc.

A descriptive linear discriminant analysis was run on the nasalance scores, which were then classified using predicative discriminant analysis. With the within-subject design, there were eleven participants in each group. This was at least five times the number of predictors (Burns & Burns, 2009) and so we proceeded with both stimuli as predictors. In a last step, the discriminant functions derived were then applied to pre-existing data sets by Karnell et al. (2004) and Bressmann et al. (2006). It should be noted that this was a somewhat unusual application of the linear discriminant analysis, as the speakers were not independent. However, as nasalance scores for mixed nasality are lacking, a within-subject design allowed for the experimental generation of such data while controlling for individual differences.
3.3 Results

The means and standard deviations for the nasalance scores and the nasalance distance are shown in Table 3.2.

Table 3.2 Mean nasalance scores and nasalance distance with standard deviations for the Zoo sentence (oral) and the first Nasal sentence as measured with the Nasometer 6450 by condition (normal, hyponasal low, hyponasal high, hypernasal, mixed low and mixed high).

<table>
<thead>
<tr>
<th>Condition</th>
<th>Zoo (oral) M</th>
<th>Zoo (oral) SD</th>
<th>Mama (nasal) M</th>
<th>Mama (nasal) SD</th>
<th>Distance M</th>
<th>Distance SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>10.2</td>
<td>3.2</td>
<td>62.0</td>
<td>4.2</td>
<td>51.7</td>
<td>3.8</td>
</tr>
<tr>
<td>Hypo Low</td>
<td>5.0</td>
<td>1.3</td>
<td>37.2</td>
<td>12.3</td>
<td>32.3</td>
<td>11.6</td>
</tr>
<tr>
<td>Hypo High</td>
<td>5.8</td>
<td>1.3</td>
<td>44.6</td>
<td>11.3</td>
<td>38.8</td>
<td>10.9</td>
</tr>
<tr>
<td>Hyper</td>
<td>53.2</td>
<td>20.2</td>
<td>70.5</td>
<td>9.9</td>
<td>17.3</td>
<td>14.8</td>
</tr>
<tr>
<td>Mixed Low</td>
<td>37.2</td>
<td>14.4</td>
<td>44.6</td>
<td>17.1</td>
<td>7.3</td>
<td>10.2</td>
</tr>
<tr>
<td>Mixed High</td>
<td>50.3</td>
<td>18.9</td>
<td>59.3</td>
<td>17.7</td>
<td>9.0</td>
<td>10.7</td>
</tr>
</tbody>
</table>
3.3.1 Repeated Measures ANOVA

A repeated measures ANOVA was run for the six conditions, two stimuli and two repetitions. Where sphericity was violated, the Greenhouse-Geisser correction was used. There were significant effects for condition \((F(2.902, 29.021) = 30.633, p < .001)\), stimuli \((F(1,10) = 139.742, p < .001)\) and a condition-stimuli interaction effect \((F(5,50) = 50.187, p < .001)\). There was no significant effect for repetition.

Post-hoc paired t-tests for the main effect of condition demonstrated that mean nasalance scores increased significantly \((p < .01)\) from hyponasal low (21.1, SD 18.5), to hyponasal high (25.2, SD 21.2), to normal (36.1, SD 26.4) and mixed low (40.9, SD 16.1). There was no significant difference between normal and mixed low \((p = .273)\). The highest mean nasalance scores were found for mixed high 54.8 (SD 18.1) and hypernasal 61.8 (SD 18.0). These two conditions did not differ from each other \((p = .181)\) but both were significantly higher than all other conditions \((p < .01)\). A paired t-test for the main effect of stimulus found the mean nasalance score of the oral sentence (26.9, SD 24.2) significantly lower than the mean nasalance score for the nasal sentence (53.0, SD 17.3; \(p < .001\)).

The nasalance score means of the condition-stimulus interaction effect are presented in Table 3.2. Post hoc paired t-tests demonstrated that, for the oral stimulus, all the conditions were significantly different from each other \((p < .01)\) except the hypernasal and mixed high condition. The mean scores increased from hyponasal low to hyponasal high to normal to mixed low to mixed high and hypernasal. The mean nasalance scores for the nasal stimulus were not significantly different for the normal and mixed high condition or the hyponasal high and mixed
low conditions, but all remaining combinations of conditions were significantly different from each other (p < .01). The lowest mean for the nasal stimulus was hyponasal low, followed by hyponasal high and mixed low, which were equivalent. At the higher end were normal and mixed high, whose means did not differ significantly, while the hypernasal condition produced the highest mean.

3.3.2 One-Way ANOVA for Nasalance Distance

A one-way ANOVA was run for the nasalance distance across the six speaking conditions. The effect of condition was significant, F(5,126) = 59.007, p < .001. Paired t-tests revealed that there were no significant differences between the means of the hyponasal high/low or the mixed high/low conditions. All remaining nasalance distance means decreased significantly from normal to hyponasal high / low to hypernasal to mixed high / low (p < .05 after REGWQ adjustment). The mean values and their standard deviations can be found in Table 3.2.

3.3.3 Discriminant Analysis

In order to determine how the six conditions differed with respect to their nasalance scores for the two stimuli and to derive a classification formula, both descriptive and predictive discriminant analysis were conducted. The oral and nasal stimuli were entered as predictor variables and the six simulated resonance conditions were the classification variables. As equal variances could not be assumed, (Box’s M p < .001) the separate groups covariance matrix option was selected (Green, Salkind, & Akey, 2000). Two discriminant functions were calculated, with a combined Wilks’lambda $\Lambda=.134$ and $\chi^2 (10) = 255.110$, p < .001. After removal of the first function, there was still highly significant discriminating power in the residual Wilks’lambda $\Lambda=.573$, $\chi^2 (4) =$
70.786, \( p < .001 \). This test indicated that the nasalance scores differentiated significantly among the conditions after partitioning out the effects of the first discriminant function. As both tests were significant, both discriminant functions were interpreted. The differences among the six conditions accounted for 76.6% of the variance in values of the first two discriminant functions and 42.8% of the variance in values of the second discriminant function.

Table 3.3 Canonical discriminant function coefficients derived from two predictors (oral and nasal stimuli) and six simulated conditions (normal, hyponasal low, hyponasal high, hypernasal, mixed low and mixed high).

<table>
<thead>
<tr>
<th></th>
<th>Function 1</th>
<th>Function 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral (Abbrv Zoo)</td>
<td>0.100</td>
<td>-0.021</td>
</tr>
<tr>
<td>Nasal (Mama)</td>
<td>-0.048</td>
<td>0.089</td>
</tr>
<tr>
<td>(Constant)</td>
<td>-0.149</td>
<td>-4.162</td>
</tr>
</tbody>
</table>

The canonical discriminant function coefficients are displayed in table 3.3. Discriminant function 1 gave greater weight to the oral stimulus maximally separating the presence of hypernasality from the absence of hypernasality. The second function gave greater weight to the nasal stimulus, maximally separating the presence of hyponasality from the absence of hyponasality. As per table 3.3 the discriminant function formulas were

\[
D_1 = (.100)\text{Oral} - (.048)\text{Nasal} - .149
\]

\[
D_2 = (-.021)\text{Oral} + (.089)\text{Nasal} - 4.162
\]
Table 3.4 Function values of group centroids for six simulated conditions (normal, hyponasal low, hyponasal high, hypernasal, mixed low and mixed high)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Function 1</th>
<th>Function 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>-2.098</td>
<td>1.151</td>
</tr>
<tr>
<td>Hyponasal Low</td>
<td>-1.439</td>
<td>-0.945</td>
</tr>
<tr>
<td>Hyponasal High</td>
<td>-1.706</td>
<td>-0.306</td>
</tr>
<tr>
<td>Hypernasal</td>
<td>1.782</td>
<td>1.005</td>
</tr>
<tr>
<td>Mixed Low</td>
<td>1.433</td>
<td>-0.974</td>
</tr>
<tr>
<td>Mixed High</td>
<td>2.028</td>
<td>0.069</td>
</tr>
</tbody>
</table>

Each participant’s set of scores produced a pair of function values. The minimal Mahalanobis distance between that pair and those of the condition centroids (displayed in table 3.4) determines which condition the set of scores is predicted to belong to. When the predicative formulas were applied to the participants’ nasalance scores, 64.4% of the conditions were correctly classified. Of the 22 pairs of nasalance scores for the normal condition, three were classified as hypernasal. For the hyponasal low condition, two were classified as normal and five were classified as hyponasal high. In the hyponasal high condition, three were classified as normal and six were classified as hyponasal low. The hypernasal condition saw two tokens classified as normal and five classified as mixed high. The mixed low condition had four classified as hypernasal and two as mixed high. Finally, the mixed high condition had ten classified as hypernasal and five classified as mixed low.
From a clinical point of view, hyponasality only becomes relevant if it is obstructive and nasal patency is reduced. In the hyponasal high and mixed high conditions, nasal patency was less compromised. For a discriminant analysis, the attributes used to separate the groups are meant to discriminate quite clearly between the groups with minimal overlap between the categories (Burns & Burns, 2009). Therefore, these two conditions were removed and the linear discriminant analysis was repeated with the remaining four speaking conditions.

Two significant discriminant functions were calculated, with a combined Wilks’lambda $\Lambda = .111$ and $\chi^2(6) = 184.993$, (p < .001) and a residual Wilks’lambda $\Lambda = .439$, $\chi^2 (2) = 69.129$, (p < .001). The differences among the four conditions accounted for 74.8% of the variance in scores of the first two discriminant functions and 56.1% of the variance in scores of the second discriminant function.

Table 3.5 Canonical Discriminant Function Coefficients derived from two predictors (oral and nasal stimuli) and four simulated conditions (normal, hyponasal low, hypernasal mixed low).

<table>
<thead>
<tr>
<th></th>
<th>Function 1</th>
<th>Function 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral (Abbrv Zoo)</td>
<td>0.990</td>
<td>-0.016</td>
</tr>
<tr>
<td>Nasal (Mama)</td>
<td>-0.050</td>
<td>0.094</td>
</tr>
<tr>
<td>(Constant)</td>
<td>0.068</td>
<td>-4.594</td>
</tr>
</tbody>
</table>

The canonical discriminant function coefficients appear in table 3.5, providing the discriminant function formulas
D1 = (.99)Oral − (.050)Nasal + .068
D2 = (-.016)Oral + (.094)Nasal − 4.594

Table 3.6 displays the function values for the four condition centroids. When the predictive formulas were applied to the participants’ nasalance scores in the four conditions, 88.6% of the simulations were correctly classified. For the normal condition, one of the 22 sets of scores was misclassified as hyponasal. Two tokens in the hyponasal condition were misclassified as normal. Two scores from the hypernasal condition were also misclassified as normal, while one was misclassified as mixed. Finally, four scores in the mixed condition were misclassified as hypernasal.

Table 3.6 Function values of group centroids for four simulated conditions (normal, hyponasal low, hypernasal and mixed low).

<table>
<thead>
<tr>
<th>Condition</th>
<th>Function 1</th>
<th>Function 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>-2.02</td>
<td>1.041</td>
</tr>
<tr>
<td>Hyponasal Low</td>
<td>-1.305</td>
<td>-1.189</td>
</tr>
<tr>
<td>Hypernasal</td>
<td>1.802</td>
<td>1.162</td>
</tr>
<tr>
<td>Mixed Low</td>
<td>1.522</td>
<td>-1.014</td>
</tr>
</tbody>
</table>
Figure 3.1 Scatterplot of function values with group centroids (1 = normal, 2 = hyponasal low, 3 = hypernasal, 4 = mixed nasality)

3.3.4 Application of Discriminant Analysis

Next there was a retrospective application of the linear discriminant classification algorithm to two data sets. Using the SPSS wizard function, the discriminant functions from the four-condition model were applied to data from Bressmann et al. (2006). The data base included nasalance scores and perceptual judgements of hypernasality. Their stimuli were the full Zoo Passage and all Nasal Sentences. The discriminant functions classified 27 of the 50 control
participants as normal, 16 as hypernasal and seven as hyponasal. Of the eight participants in the mild hypernasality group, two were classified as normal, one as hyponasal and five as mixed nasality. Four of the eleven participants of the moderate hypernasality group were classified as hypernasal and seven as mixed nasality.

The discriminant function was also applied to a set of nasalance scores that were published by Karnell et al. (2004). In this study, the authors investigated the effect of palatal lift appliances on nasalance scores. The participants had a variety of causes of velopharyngeal dysfunction. Only one patient had velopharyngeal dysfunction related to cleft palate. A typo was noted in the data table by Karnell et al. (2004) but the erroneous value could be deduced from the reported mean. The study also provided the individual perceptual scores of hypernasality and hyponasality from normal to severe on scales from one to six. For the purpose of this classification exercise, a rating above one, for both hypernasality and hyponasality, was taken as a perception of mixed nasality. Without the palatal lift appliance, two participants were perceived as having normal nasalance, one had mixed nasalance and the remaining sixteen were hypernasal. With their palatal lift appliances, five were perceived as having normal resonance, eight as hyponasal, two as mixed and four as hypernasal. The predictions derived from the discriminant functions of the nasalance scores matched 15 of the 19 pairs of perceptual assessments when the palatal lift were not in place. A participant perceived to have normal resonance and one participant perceived to have mixed resonance were predicted to be hypernasal. One of the hypernasal participants was predicted to be normal and another was predicted to have mixed nasality. When the palatal lifts were in place, the predictions matched eight of the 19 pairs of perceptual assessments. The mismatches included four participants perceived to have normal resonance where two were predicted to have hypernasality, one was predicted to be hyponasal and one was predicted to have
mixed nasality. Of the five speakers perceived to have hyponasal resonance, two were predicted to have normal resonance and three were predicted to have mixed resonance. Finally, a participant perceived to have mixed resonance with the palatal lift in place was predicted to have hypernasal resonance.

3.4 Discussion

The goal of the present study was to develop a tentative classification formula for different types of resonance disorders. The data for this experiment were obtained from speakers who provided normal samples as well as simulations of different resonance disorders. The descriptive statistics for the results indicated that the simulations were overall successful and appeared reasonably similar to data that would typically be obtained from clinical participants.

A repeated measures ANOVA had a highly significant text-condition interaction effect and revealed significant differences in scores across the conditions. A series of post hoc paired t-tests confirmed that the participants were able to produce a wide range of significantly different nasalance scores. Where there were expectations, the changes in the nasalance scores conformed to those expectations. They decreased when nasal obstruction was applied and increased when a hypernasal resonance was used. As for the mixed nasalance conditions, the mean nasalance score of both the mixed high and mixed low conditions were higher than the normal conditions for the oral stimulus. The mixed low condition was significantly lower than the normal condition for the nasal stimulus, but the mixed high condition was not.
The mean for the oral stimulus in the normal condition (10.3, SD 3.2) was lower and less variable than previously found among normal speakers of Southern Ontario. Using the Nasometer 6200 and similar or identical stimuli, reported values have been 12 (SD 6) (Seaver et al., 1991), 13.45 (SD 5.90) and 11.62 (SD 4.33) (Bressmann, 2005). For the same Nasometer model 6450 and stimulus, the mean was 13.12 (SD 6.35) (Chapter 2). The mean for the nasal stimulus in the normal condition, 62.0 (SD 4.2) was comparable to, but less variable than in, other studies with similar speakers. Previously reported means for nasal stimuli with the Nasometer 6200 were 61 (SD 7) (Seaver et al., 1991), 57.90 (SD 6.69) and 57.01 (SD 7.64) (Bressmann, 2005). Using the Nasometer 6450, the same nasal sentence was found to have a mean of 63.54 (SD 6.27) (Chapter 2). The means for the nasal stimulus in the hyponasal conditions were both well below the 50% cut off score proposed by Dalson et al. (1991b). Likewise, the mean nasalance score for the oral stimulus in the hypernasal condition surpassed the proposed cut offs of 28% (Dalston et al., 1993) and 32% (Hardin et al., 1992). In fact, the mean nasalance for the hypernasal condition nearly matches the mean nasalance of 53 (SD 7.2) of a group perceived to have severe hypernasality (Dalston et al., 1993).

The simulation of different resonant disorders by the normal speakers led to very clear-cut quantitative results. However, this simulation produced little of the overlap in nasalance scores between normal and mildly hypernasal resonance for the oral stimulus that can affect the exact calculation of cutoff scores in clinical research (Dalston et al., 1993; Bressmann et al., 2006). On one hand, this may have been due to a relatively small range of scores produced by the speakers in this study for the normal condition compared to the variability observed in previous research with larger groups (Bressmann, 2005; Gildersleeve-Neumann & Dalston, 2001; Seaver et al., 1991). On the other hand, the participants in the present study were instructed to produce only
severe hypernasality. Perceptually, both investigators had the impression that the participants were successful at this task, and the nasalance scores for the hypernasal condition confirmed this impression. It was the purpose of the present study to use simulations to create prototypical nasalance profiles for different resonance disorders. In future clinical research it would be important to find out in how far these simulated profiles differed from actual clinical data.

The one-way ANOVA for the nasalance distance was highly significant, indicating that the nasalance distances differed by condition. The mean nasalance distance of the normal condition was significantly greater than that of the hypernasal condition, as expected from previous research (Bressmann et al., 2000, 2006). The present research demonstrated that hyponasal and mixed resonance disorders should also have significantly shorter nasalance distances. While the nasalance distance in hypernasality is shortened because of elevated scores for oral stimuli, the nasalance distance for hyponasality is shortened because of lower scores for nasal stimuli. The nasalance distance for the mixed nasality conditions were shortened by elevated scores for the oral stimulus, and for the mixed low condition only, lowered scores for nasal stimulus. The nasalance distance might serve as a useful supplement to the mean nasalance scores but more research is needed. While the ANOVA showed significant differences between the disordered conditions, the potential discriminative value of the nasalance distance was not evaluated in the present study. As both the nasal and oral stimuli were entered into the discriminant function, the addition of the nasalance distance would have added redundant information.

The linear discriminant analysis classified 64.4% of the data correctly into the six conditions, based on the nasalance scores. This result was better than chance alone (chance level 16.7%). By removing the hyponasal high and mixed high conditions, where nasal obstruction had less impact
on nasalance scores, a better rate of 88.6% accurate predictions was achieved, well above the 25% chance level. Some of this increase in classification accuracy was due to the decreased number of categories. In fact, the ratio of the predictions to chance decreased from 3.9 to 3.5 with the decrease in categories. However, this reduced the overlap in values between categories, thus allowing us to derive two first tentative formulas for a classification of resonance disorders.

The tentative application of these discriminant analysis functions to pre-existing data sets, which included nasalance scores and perceptual ratings, were moderately successful. When applied to the data set by Bressmann et al. (2006), the discriminant functions could only corroborate normal resonance and the presence of hypernasality (alone or mixed - with hyponasality) as there were no perceptual measures of hyponasality for the data set. Of the nineteen participants with a cleft palate and some perceived degree of hypernasality, 21% were classified as hypernasal and 63% were reclassified as having mixed nasality. There were three disagreements between the perceptual judgements and the predictions. They represented 16% of the speakers with cleft palate, all of which were in the mildly hypernasal group. Interestingly, the corroboration of normal resonance in the control speakers proved even more difficult. Nearly half of the normal speakers were classified as hypernasal (32%) or hyponasal (14%). This may be due, in part, to the aforementioned limited range of scores the participants of the simulation experiment produced for the oral stimulus in the normal condition. When the groups were regrouped into a combined hypernasal/ mixed nasality and a combined normal/ hyponasal group to mirror the perceptual classifications by Bressmann et al. (2006), the sensitivity of the discriminant functions for predicting the presence of hypernasality (alone or mixed) was .84 and the specificity (to exclude 27 normal & 7 hyponasal) was .68 for an overall efficiency of .72. These results are presented in Table 3.7. The sensitivity compared well with those found by Dalston et al. (1993) and Hardin
(1992) but the specificity was lower. The specificity was also lower than that found by Bressmann et al. (2006) using the nasalance distance. However, the sensitivity was comparable (Bressmann et al., 2006).

**Table 3.7 – Results for sensitivity and specificity of the discriminant functions when applied to the data set by Bressmann et al. (2006). Only presence or absence of hypernasality were evaluated.**

<table>
<thead>
<tr>
<th>Discriminant Prediction</th>
<th>Hypernasality (perception)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
</tr>
<tr>
<td>Hypernasal / Mixed Nasality</td>
<td>16</td>
</tr>
<tr>
<td>Normal / Hyponasal</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.84</td>
<td>0.68</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Some of the classifications of normal speakers as hypernasal or hyponasal may be attributed to differences in the stimuli and Nasometer models used. Full versions of the oral and nasal stimuli were spoken in Bressmann et al. (2006) while an abbreviated version was used by the participants of the present study. The full version of the Zoo Passage has been found to score two points higher, on average, than the abbreviated form (Bressmann, 2005). In addition, the Nasometer 6200 used in Bressmann et al (2006) has been found to score two points lower for the nasal stimulus than the Nasometer 6450 (see Chapter 2).
The overall agreement between Karnell et al.’s (2004) perceptual evaluations and the predictions of the discriminant functions was 61%. While the agreement is not very high, it is much higher than chance alone. However, the concordance was much greater without (79%) than with the palatal lift prostheses (42%). It should be noted that Karnel et al. (2004) used the Nasometer 6200 with a shortened version of the Zoo Passage and the full set of Nasal Sentences. The differences in stimuli and instruments may account for some of the discrepancies between the perceptual ratings and the predictions based on the discriminant functions.

A caveat for the interpretation of the study findings was the small sample size. Finding individuals who could successfully simulate hypernasality proved challenging. The simulations of hypernasality were also predominantly severe, leading to higher scores than would be found in many typical clinical populations. Although small, the sample size was large enough to show that the nasalance scores of mixed nasality differ from hyponasality and hypernasality. More research with clinical populations will be needed to corroborate the validity of the derived classification formulas.

### 3.5 Conclusion

The present study used linear discriminant analysis as a tentative step towards an assessment of resonance disorders. The simulations across the conditions were distinct enough from each other to be used to derive a formula that predicted resonance above chance level. The initial results are promising and we see potential in this approach. The validity of the formula would be improved with increased sample size. Increasing the number of normal condition scores would be quite
straightforward, yet future hyponasal, hypernasal and mixed scores should come from a clinical population to reflect the range of severities.
4. General Conclusion

Nasometry aids in the diagnosis and treatment of resonance disorders by providing a quantitative measure of the oral-nasal resonance balance to supplement the clinical judgement. However, the instrument is not perfect and the listener-Nasometer agreement is sometimes unsatisfactory. The present study first compared a new Nasometer, model 6450, to the performance of a previously vetted model, the Nasometer 6200. There was also a gap in knowledge regarding the day-to-day variability of nasalance scores inherent to a nasal stimulus. The information was required to determine the variation in scores introduced by the new machine and the speech stimuli. The newer Nasometer 6450 was then used to attempt to improve the diagnostic procedures for resonance disorders.

When the Nasometers were compared using synthetic stimuli, they provided similar scores across a range of frequencies. The nasalance scores generated from various stereo panoramas deviated from what was expected in theory. The Nasometer’s sound separator plate is meant to prevent up to 25 dB SPL from reaching the opposite microphone. However, it was not known how this would affect the nasalance readings. For both instruments, the amount of sound energy that actually spills over to the opposite side was found to be 7 nasalance points.

Next, the Nasometers were compared using normal participants. The Nasometer 6450 scored one point higher than the Nasometer 6200 for a balanced stimulus and two points higher for a nasal stimulus. No difference was found when the stimulus was oral (void of nasal sounds). For both Nasometers, the day-to-day variability ranged from six points for an oral stimulus to eight points for a nasal stimulus. As the day-to-day variability is substantially greater than the difference
between the instruments, acquiring a new Nasometer is not likely to impact clinical practice. It is important for a clinician to understand that the nasalance score is less precise than its presentation makes it appear. However, it is helpful to understand the normal variability of scores when monitoring a patient’s speech over time.

The third chapter of the thesis focused on the development of an assessment protocol based on nasalance scores. The research to date has focused primarily on hypernasality and all but excluded mixed nasality. A fuller picture of resonance disorders and the scores associated with them is required. This diagnostic improvement is also needed because inter-listener, intra-listener and listener-Nasometer agreement have been problematic. Using the Nasometer 6450 and simulations of disordered resonance, nasalance scores of normal, hyponasal, hypernasal and mixed nasality resonance were collected. The results showed that mixed nasality produced a range of scores that is different from hypernasality and hyponasality. Once significant differences in nasalance scores between the conditions were confirmed, they were used in a linear discriminant analysis to create a preliminary diagnostic formula. The formula was applied to pre-existing data sets of nasalance scores. The resulting predictions of the accompanying resonance disorder (as judged by listeners) were reasonable. After further refinement of the formulas, the nasometer could point the clinician towards the general type of the patient’s resonance disorder. However, the current formulas are hampered by the differences in scores between the Nasometer 6450 (used in creating the formula in Chapter 3) and the Nasometer 6200 (used in the pre-existing data sets). More importantly, the normal participants in our study had a lower mean for the oral stimulus and a smaller range of scores for both stimuli in the normal condition than found in the study used for comparison (Bressmann et al., 2006). This led to a relatively small range of normal scores and a formula biased towards disordered resonance.
To improve the formula, it would be desirable to increase the number of speakers for each condition to at least 25. The data for disordered resonance should come from clinical populations in order to capture the full range of their scores. Ideally, an instrumental measure of nasal patency such as rhinomanometry (Williams et al., 1990) would complement the nasalance scores. Nevertheless, additional variables such as nasal emission (Dalston et al., 1991b), a pharyngeal flap (Hardin et al., 1992) or a maxillofacial prosthesis (Karnell et al., 2004) can all add measurement noise to the nasalance score.

Eventually, once enough data points have been collected for each resonance disorder in a clinical setting, a diagnostic chart could be created. In such a chart, a clinician could relate the patient’s nasalance score for an oral stimulus to their score for a nasal stimulus and locate the predicted resonance based on those scores. A hypothetical illustration of such a chart appears in figure 4.1. Test-retest variability could be documented graphically in the chart (by marking squares).

Alternately, a simple software could be developed whereby the scores from the oral and nasal stimuli are entered. A prediction of the most likely resonance would be provided, along with its probability of being correct. Where the probability is less than one, a prediction for the second most likely resonance (and its probability) would be provided. The software could include correction factors for specific Nasometer models and stimuli used. By making the nasometric assessment less exact, we might it make it more robust and reliable.
Figure 4.1 Hypothetical resonance matrix for normal, hyponasal, hypernasal and mixed nasality based on scores from the Zoo Passage (oral) and the Nasal Sentences using the Nasometer 6450.

In the end, the issue of listener agreement is unlikely to ever be fully resolved because, ultimately, the ear is subjective. The Nasometer will faithfully measure what is presented to it but it is important that we ask the right questions of the instrument. Rather than striving for more accurate measurements and continuing to chase exacting cut-off values, the clinician could consider allowing for more fuzziness in their interpretation of nasometric scores. In this way of thinking, the nasometer could be used like a compass, pointing the clinician towards the general category of the patient’s resonance disorder.
5. References


Watterson, T., Lewis, K., & Brancamp, T. (2005). Comparison of nasalance scores obtained with
the Nasometer 6200 and the Nasometer II 6400. *Cleft Palate Craniofacial Journal, 42,*
574-579.

Watterson, T., & Lewis, K. E. (2006). Test-retest nasalance score variability in hypernasal
speakers. *Cleft Palate Craniofacial Journal, 43,* 415-419.

Craniofacial Journal, 38,* 119-125.


Impairement. In M. J. Ball, M. R. Perkins, N. Müller & S. Howard (Eds.). *The Handbook of


_Comprehensive Cleft Care_ (pp. 627-640). New York: McGraw Hill Medical.