THE PERCEPTUAL AND INSTRUMENTAL ASSESSMENT OF NASALITY AND NASAL AIRFLOW ERRORS ASSOCIATED WITH VELOPHARYNGEAL DYSFUNCTION

Caitriona Sweeney, M.Sc., M.R.C.S.L.T.

A dissertation submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the School of Clinical Speech and Language Studies, University of Dublin, Trinity College.

July 2000

DECLARATION

I hereby declare that this thesis is entirely the work of the undersigned. It has not been submitted for a degree at any other University.

I agree that the depository library may lend or copy this thesis, in whole or in part, without further reference to the author.

Signed: Caitriona Sweeney

Caitriona Sweeney

DEDICATION

This work is dedicated to Sarah, Michael, Graham, Yvonne, John, Hannah, Jonathan, Holly, Lorcan, Joanna, Adam, Emmet and all the children and families with whom I have had the pleasure to work. They have been an inspiration.

Acknowledgement

The author wishes to express her deep appreciation to Dr. Debbie Sell for her guidance, supervision, hard work and wonderful support in her role as supervisor and colleague. Sincere thanks also to Professor Jim Malone for his guidance, persistent support and friendship throughout the last five years. Many thanks also to Professor Pam Grunwell and Margaret Leahy for their help, particularly in the initial stages of this thesis.

Sincere gratitude also to Mr. Fran Hegarty, Department of Medical Physics and Bio Engineering, St. James Hospital, for his incredible work, moral support and friendship throughout the early stages in setting up in the instrumentation. Without his help this thesis could not have been completed. My gratitude also to Myra O'Regan, Department of Statistics, TCD, for all her help, patience and guidance with the statistics, and her continued positive encouragement.

Many thanks to Brian Sweeney for all his help with the 'IT' involved in putting this thesis together, and for constantly being at the end of the phone to solve any computer problems. To Charlie and Colette Sweeney, I wish to express my deepest gratitude for their role as parents, providing constant support, and for their role as grandparents, providing "child collection and minding" services. My only regret is that Charlie did not see the final result. Also thanks to Orna for all her support, especially during this last difficult year.

To my colleagues in the Speech and Language Therapy Department of The Children's' Hospital, Temple Street, sincere gratitude. To Eilis Murphy, for her constant support and help throughout the entire research process, to Mary Pat O'Malley and Siobhan Ryan for all their help.

To all the children and families of the Ranelagh Multi Denominational School and the Cleft Palate Unit at Temple Street for their help with data collection. Also to Winifred

Ryan for her help with normative data. This thesis could not have been undertaken without the financial support of The Child Health Foundation, Temple Street.

Finally, my deep appreciation to my husband Mondy for his support, which ranged from moral support, child-minding and reading to providing me the time to complete this task. To my children Fionan, Eoin and Oisin for allowing me to take over the computer and for the wonderful diversion they provided for me during the last few years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This study aimed to develop a reliable and valid perceptual profile for the assessment of nasality and nasal airflow errors in speech. To date, clinicians rely on simple categorical or numerical scales of nasality and nasal airflow errors to assess speech. Such scales provide limited information on the type or degree of the presenting speech problems. This new scale aimed to describe the nature and the degree of these characteristics, thus improving reliability and validity. The second aim of the study was to assess the validity of the Perceptual Profile using instrumental assessment for investigations of nasality and nasal airflow errors and thus develop a protocol for assessment of nasality and nasal airflow errors in speech.

In order to develop the Perceptual Profile, working definitions of the terms used to describe nasality and nasal airflow errors were developed. The definitions formed the basis of the Perceptual Profile and qualitative descriptions of error categories were devised. The resulting scale was then tested rigorously for reliability, using percentage of agreement and kappa scores. Intra-rater reliability was assessed by rating the speech of twelve children presenting with nasality and/or nasal airflow errors twice from audio tape recordings. Results indicated good to excellent intra-rater reliability of the Perceptual Profile. Inter-rater reliability was assessed by comparing ratings of three Speech and Language Therapists (one specialist, one experienced and one inexperienced in the area). The speech samples of 20 children were analysed; percentage agreement and kappa scores were calculated for each pair of raters. Results indicated good inter-rater agreement for perceptual ratings of nasality and nasal airflow errors. However, kappa results were variable, ranging from good to poor. Agreements of ratings in the present study compared well with previous studies. The use of kappa analysis in the evaluation of reliability of nasality and nasal airflow assessment is new and comparison of the kappa analysis of the perceptual profile with other scales is not possible. The variable findings of reliability of the perceptual assessment demonstrates the need to supplement perceptual assessments with instrumental measurements.

Perceptual ratings of nasality were compared to instrumental measurements of nasality using the Nasometer, while perceptual ratings of nasal airflow errors were compared to

instrumental measurements of flow and pressure using the PERCI SARS system. Normative data from instrumental measurements is required before instrumentation can be used to assess pathological speech. Normal pressure/flow measurements from 152 normal speaking children were obtained by Zajac (1998, personal communication). Previous studies have indicated the need to obtain local normative data for the Nasometer. In this study, seventy normal English-speaking Irish children were assessed on the Nasometer using a novel speech stimulus. The speech stimuli available to date have been found to be difficult for children (Watterson, Hinton & McFarlane, 1996) and the present study found that the American stimuli were culturally biased. The speech stimulus used for the present study allowed for nasometric analysis of sentences, which were categorised according to consonant type. The normative data obtained in the study will now provide baseline normal data for use of the Nasometer in Ireland.

Fifty children presenting with nasality and/or nasal airflow errors were assessed using the perceptual scale, the Nasometer and the pressure/flow system. Results were compared using correlational analysis, test sensitivity, specificity and overall efficiency ratings. Results indicated a strong relationship between perceptual ratings of nasality and nasalance scores on specific speech stimuli. There was a strong relationship between perceptual ratings of nasal emission and pressure/flow measurements, particularly nasal flow and velopharyngeal port area measurements. A weak relationship was found between pressure/flow measurements and perceptual ratings of the following nasal airflow errors: nasal turbulence, nasal fricatives and velopharyngeal fricatives.

Reliability and validity results identified sections of the Perceptual Profile that required revision. Overall results indicate that the Perceptual Profile has variable reliability, with acceptable agreement and, is a valid tool for the assessment of nasality and nasal airflow errors in the speech of children with velopharyngeal dysfunction. Results also indicate that the perceptual and instrumental assessment protocol is a valid and reliable assessment, which could be used in specialist centres for management of children with cleft palate and velopharyngeal dysfunction.

CHAPTER 1

INTRODUCTION

Disorders of nasal resonance and articulation are usually the primary speech problems in cases of cleft palate, velopharyngeal dysfunction, or nasal obstruction (McWilliams, Morris & Shelton 1990; Bzoch, 1989). Sell (1999) reported that nasality and nasal airflow errors are considered primary outcome measures of palatal surgery. Hence, the importance of reliable and valid assessment of these speech parameters is underscored. Assessment of nasality and nasal airflow constitutes an important aspect of a comprehensive assessment of speech of individuals with repaired cleft palate and/or velopharyngeal dysfunction (Glossary, Appendix 1). It not only serves to evaluate the speech status of such individuals, but indirectly evaluates velopharyngeal function. Because problems of nasality and/or nasal airflow are often indicative of a problem with velopharyngeal function, some of the approaches to assessment of the speech problems are referred to as indirect measures of velopharyngeal function (Moon, 1993; Shprintzen, 1995).

Recent literature provides descriptions of articulation disorders associated with cleft palate and velopharyngeal insufficiency (Trost, 1981; Eurocleft Speech Group, 1993; 1993; Sell, Harding & Grunwell, 1994; 1999). Much attention has been paid to the investigation of articulatory errors associated with cleft palate and velopharyngeal dysfunction; however, there has been limited investigation into classification and assessment of nasality and nasal airflow errors (Grunwell & Harding, 1996). To date perceptual assessment of nasality has taken the form of equidistant rating scales (e.g. equidistant scales from 1 to 5, where 1 indicates no abnormality and 5 indicates severe abnormality) or a severity scale (i.e. designating the degree of severity in terms of mild/moderate/severe). These scales, however, do not comprehensively define all the points on the scale in terms of what is perceived by the listener. This lack of definition has resulted in poor reproducibility of assessment and weak inter-judge reliability (Wirz & Mackenzie Beck, 1995). More recent approaches to assessment of nasality have begun to define scalar points. For example, the Great Ormond Street Speech Assessment

(GOS.SP.ASS) (Sell, Harding & Grunwell, 1994) describes the parameters of hypernasality and hyponasality on a 3 point scale, where 0 indicates no nasality and 1 and 2 are defined according to the severity of nasality. However, this scale has limited descriptions of severity. In the revised GOS.SP.ASS 1998, the scale is expanded to 4 points, and more detailed definitions are provided. In the Cleft Audit Protocol for Speech (CAPS), (Harding, Harland & Razzell, 1997) details of consistency are included in the severity rating. For example, a rating of 1 indicates mild consistent hypernasality, whereas a rating of 2 indicates moderate consistent hypernasality. Such a scale is not comprehensive. It does not, for example, allow for a rating of moderate inconsistent hypernasality.

Nasal airflow errors have been assessed perceptually in a similar manner, resulting in similar problems. Furthermore, some authors do not distinguish between the different types of nasal airflow errors. For example, nasal turbulence has been described as a more severe form of nasal emission (McWilliams et al., 1990; Stengelhofen, 1990) (Glossary, Appendix 1). Others recently describe nasal turbulence as a separate feature (Sell et al., 1994), reporting a smaller velopharyngeal port gap during the production of nasal turbulence than during the production of nasal emission (Kummer, Curtis, Wiggs, Lee and Strife, 1992). Furthermore, McWilliams et al. (1990) describe two types of nasal turbulence. Intra-nasal turbulence is produced within the nasal cavity, whereas nasal turbulence is produced at the velopharyngeal sphincter. The different approaches to assessment of nasal airflow errors has led to confusion regarding the speech problems of patients with cleft palate, velopharyngeal dysfunction or nasal obstruction, and a lack of reliable and valid assessment scales.

The GOS.SP.ASS addresses some of the issues of perceptual assessment of nasality and nasal airflow; however, this is a screening assessment which has not been validated using instrumentation. The CAPS assessment is used for audit purposes, it has insufficient detail for clinical diagnoses, and it has also not been validated. Information on consistency and frequency of nasal airflow errors should be included in assessment of nasal airflow in speech (Shprintzen, 1995). Details of frequency of nasal airflow errors are not included in the GOS.SP.ASS or the CAPS. The Perceptual Profile developed for this study aims to provide a comprehensive assessment of nasality and nasal airflow errors, for example, describing nasal airflow errors in terms of strength, frequency and consistency.

In addition to the problems of the types of scales used for the perceptual rating of nasality and nasal airflow errors in speech, other difficulties in perceptual assessment of speech have been identified. The phonetic context of the speech sample, the presence of articulatory errors, listening conditions and listener training have been found to influence perception of nasality (McWilliams et al., 1990; Albery & Grunwell, 1993; Sell et al., 1994; 1999; Kent, 1996). Despite these difficulties, the perceptual assessment of nasality and nasal airflow is of utmost importance as the ultimate test of acceptability of speech involves the perceptual acceptability to the listener (Moll, 1964). Hence, the assessment of articulation and nasality is subjective, but instrumental measurements can compliment these evaluations.

The primary aim of this study is to develop a detailed Perceptual Profile for assessment of nasality and nasal airflow in patients with cleft palate, velopharyngeal dysfunction or nasal obstruction, and to rigorously evaluate the reliability and validity of the Perceptual Profile. A detailed profile should allow observation of any change in nasality and/or nasal airflow following therapy or surgery and in this way help evaluate treatment approaches. Planning specific therapy should be facilitated by a detailed assessment of speech, and indications for surgery can be identified. Such a profile should improve the exchange of information between speech and language therapists and other members of the team, and between speech and language therapists themselves.

The relationship between perceptual judgements of hypernasality, some instrumental measurements of nasality and nasal airflow and observation of velopharyngeal function have been examined (McWilliams, Glaser, Phillips, Lawrence, Lavorato, Berry & Skolnick, 1981; Karnell, Folkins & Morris, 1985; Dalston, Warren & Dalston, 1991). Studies have evaluated the use of the Nasometer as an instrumental assessment of nasality (Dalston, Warren & Dalston. 1991a; 1991b) (Glossary, Appendix 1). Results indicated that factors such as speech stimuli, language, dialect and gender influence the scores obtained from the Nasometer (Seaver, Dalston,. Leeper & Adams, 1991; Watterson, Lewis & Deutsch, 1998a; Karnell, 1995). The need to obtain normal nasalance scores for each dialect has been underscored (Seaver et al., 1991; Trindade, Genero & Dalston, 1997; van Doorn & Purcell, 1998). Further investigations into the relationship between perceptual measurements and Nasometric measurements and the effects of the speech stimuli on these relationships have been recommended (Watterson, McFarlane & Wright, 1993; Karnell, 1995). The present study aims to obtain normative data for Irish children using the Nasometer and to evaluate the relationship between perceptual ratings of nasality, using the Perceptual Profile and nasalance scores.

Instrumental assessments of nasal airflow errors have been developed for research, but few have proved to be reliable and valid assessment tools (Ellis, Flack, Curle & Selly, 1978; Mirlohi, Kelly & Manley, 1994). Studies have evaluated the usefulness of the pressure/flow system developed by Warren (1979) (Morr, Warren, Dalston & Smith, 1989; Warren, Dalston & Mayo, 1993b; Dalston & Warren, 1986). One study assessed the relationship between the various pressure/flow measurements (Laine, Warren, Dalston & Moor, 1988). Other studies evaluated the relationship between pressure/flow measurements and perceptual ratings of nasality (Dalston & Warren, 1986; Warren, Dalston & Mayo, 1994). However, scant attention has been paid to the relationship between perceptual evaluations of nasal airflow errors in speech and instrumental measurements of air pressure and flow. This study aims to assess the relationship between ratings for nasal airflow errors in speech on the descriptive Perceptual Profile and pressure/flow measurements.

1. 1 RESEARCH TITLE

The Perceptual and Instrumental Assessment of Nasality and Nasal Airflow Errors Associated with Velopharyngeal Dysfunction

1. 2 AIMS

1) To develop a reliable descriptive Perceptual Profile of nasality and nasal airflow errors in speech;

2) To validate the descriptive scale by comparing ratings on these scales with instrumental measurements of nasality and nasal airflow;

3) To obtain nasalance scores for normal Irish English-speaking children;

4) To develop a perceptual and instrumental protocol for the specialist assessment of nasality and nasal airflow errors in speech for children with cleft palate, velopharyngeal dysfunction or nasal obstruction.

1. 3 Outline of the Thesis

A review of the relevant literature regarding the assessment of nasality and nasal airflow in speech is presented in *Chapter 2*. A review of perceptual assessment of speech, describing the factors which influence the listener's perceptions of nasality and nasal airflow errors, is presented in *Section 1.* The use of scaling techniques in the assessment of nasality and nasal airflow errors is evaluated and an overview of previous reliability studies is summarised. *Section 2* describes the Nasometer, a computer based instrument which measures nasal and oral acoustic energy during speech. Normal studies using the Nasometer are described and critically evaluated. Clinical studies using the Nasometer include studies investigating the relationship between perceptual measurements and nasometric measurements. *Section 3* provides a background to aerodynamic assessments of speech. The pressure/flow system used in the present study is described and results of previous studies for normal and pathological speakers are presented. The relationship between pressure/flow studies and other assessments of nasality, such as perceptual assessment and Nasometry, are discussed. The final section of this chapter is a statement of the problems in the assessment of nasality and nasal airflow errors in speech.

A study of nasometric scores for normal English-speaking Irish children is presented in *Chapter 3*. This chapter outlines the methodology, results and discussions of a study which assessed 70 children with normal speech during the production of various speech samples. Results from the Irish population are compared with the results from previous studies. These normal values are used in the main study for comparisons between normal and pathological speakers using the Nasometer, thereby determining the validity of the nasality scale in the Perceptual Profile.

Chapter 4 describes the Perceptual Framework, which was developed for the study. *Section 1* provides a definition of terms, with working definitions of nasality (hypernasality, hyponasality and Cul de Sac resonance) and nasal airflow errors (nasal emission, nasal fricatives, nasal turbulence and velopharyngeal fricatives). These definitions form the basis of the Perceptual Profile. *Section 2* describes the Perceptual Profile, while the speech stimulus used is presented in *Section 3*. *Section 4* presents four different but related reliability studies which were undertaken using the Perceptual Profile. Three intra-rater reliability studies were carried out, including an evaluation of listening conditions, and evaluations of the use of anchor stimuli. The final reliability study evaluates inter-rater reliability for three speech and language therapists. Limitations of the reliability study, general discussion of reliability results, conclusions and implications are presented.

Chapter 5 presents the methodology used in the main study. Information on the participants, assessment procedures, calibration of the equipment and the analyses are included. The results of the study are presented in *Chapter 6*. The type and severity of nasality and nasal airflow errors are presented. Reliability of the instrumentation is reported. Descriptive statistics for the instrumental results, including comparison with normative data on the Nasometer and pressure/flow system and the relationship of instrumental measures with perceptual measures are presented.

The discussion of the methodology and results is presented in *Chapter 7*. The results are discussed under the headings of perceptual assessment, nasometric assessment and pressure/flow assessments. The discussion of the instrumental results includes the relationship between perceptual and instrumental measurements, the efficiency of the speech stimuli used in assessment, the efficiency of cut-off values and possible factors which influence the relationship between measurements. The conclusions of the present study are presented in *Chapter 8*, with recommendations for a clinical protocol for assessment of nasality and nasal airflow errors in speech. Recommendations for further research are outlined.

CHAPTER 2

ASSESSMENT OF NASALITY AND NASAL AIRFLOW

Introduction

Speech is recognised as one of the primary outcome measures of palatal surgery (McWilliams et al., 1990; Grunwell, Sell & Harding, 1993). In order to make decisions about "the need for modification of oral structures, the appropriateness of speech or language intervention, and the planning and execution of therapeutic procedures" an in-depth assessment of speech production is essential (McWilliams et al., 1990; p.3).

Cleft palate, non cleft velopharyngeal dysfunction and nasal obstruction have been grouped together as possible causes of the speech anomalies of resonance, nasal airflow and articulation (Glossary, Appendix 1). A cleft of the palate is an opening in the palate resulting in the continuous passage between the mouth and nose, where the palate serves as the roof of the mouth and the floor of the nose (Shprintzen, 1995). Cleft palate is a congenital condition and is the most common craniofacial anomaly. According to Watson (2000), there is an almost infinite variation in the presentation of cleft lip and palate. He stresses the difficulty in classification of clefts, where the simple classifications are required for 'everyday use', but they fail to distinguish the severity within the groups. In general the classification developed by the American Cleft Palate Association (Harkins, Berlin, Harding, Longacre, Snodgrasse, 1960) is adapted and used. This system has three main categories: cleft of prepalate (or primary palate), which includes a cleft of lip, alveolar ridge and the triangle of the palate to the incisive foramen; cleft of palate (or secondary palate), which includes the hard palate posterior to the incisive foramen and the soft palate and; cleft of the prepalate and palate. Within each category the clefts are described in terms of severity and location. Clefts of the primary palate may be complete or incomplete, and they may be unilateral, bilateral or median, while clefts of the secondary palate may be complete, incomplete or submucous (Watson, 2000).

Velopharyngeal dysfunction is a generic term which denotes any type of abnormal velopharyngeal function resulting from structural deficits, neurological disorders, faulty learning or a combination of aetiologies. Abnormal speech is often the primary symptom of velopharyngeal dysfunction. Nasal obstruction may result from enlarged adenoids, deviated nasal septum, a pharyngeal flap (Appendix 1), inadequate nasal airways and chronic catarrh. This obstruction may prohibit the normal production of nasal consonants and possibly influence nasal airflow (Wyatt, Sell, Russell, Harding & Albery, 1996).

Disorders of speech primarily associated with cleft palate, non cleft velopharyngeal dysfunction and nasal obstruction are evident at segmental and suprasegmental levels. The segmental problems occur on an isolated sound and include articulatory errors of place and manner and problems of nasal airflow. Suprasegmental problems extend over units which encompass more than one segment (Hyman, 1975) and include aspects such as stress, intonation and resonance. Resonance of the voice refers to the acoustic response of air molecules to a source of sound in the oral, nasal, and pharyngeal cavities (Borden & Harris, 1980). In this case the sound source is the vibration of the larynx. The terms, which have been used to describe nasal resonance include nasality, nasal voice, and denasal voice, all of which have been used interchangeably (Laver, 1980). These terms have been used to describe both segmental features of nasality on isolated sounds and suprasegmental features of voice quality. In the present study, the term nasality will be used to describe the perceptual suprasegmental feature associated with the acoustic response of air within the coupled oral and nasal cavity. Laver (1980) states that nasality is primarily an auditory concept and that the term covers a number of auditorily similar but not identical phenomena. In normal speech, nasality is evident on nasal consonants and adjacent vowels. Disorders of nasality are commonly associated with cleft palate, velopharyngeal dysfunction and nasal obstruction. Hypernasality has been referred to as increased nasal resonance of the voice (Wyatt et al., 1996), with the perception of inordinate nasal resonance during production of vowels due to inappropriate coupling of the oral and nasal cavities (D'Antonio & Scherer, 1995). Hyponasality has been referred to as a reduction in normal nasal resonance resulting from a blockage of the nasal airway (D'Antonio & Scherer, 1995). Detailed working definitions of nasality, hypernasality and hyponasality are presented in Chapter 4, Section 1.

Segmental errors associated with cleft palate and non cleft velopharyngeal dysfunction include errors of nasal airflow and articulatory errors. For the purpose of this study, nasal airflow errors is a generic term describing the inappropriate escape of air through the nasal cavity during production of voiced and voiceless oral pressure consonants, and is usually associated with incomplete velopharyngeal closure. Nasal airflow errors consist of the following, based on the literature:

Nasal emission is audible escape of air from the nasal cavity accompanying production of oral pressure consonants. When airflow is constricted within the nasal cavity, it is noisy and therefore audible (Grunwell & Harding, 1998). Nasal emission has a frictional but no turbulent or snorting quality. When this occurs during consonant production there is both oral and nasal release of air, as the nasal emission accompanies the sound (Sell et al., 1994).

A *nasal fricative* also has the frictional sound produced by air passing through the nasal cavity when there is incomplete velopharyngeal closure. However, there is a complete or almost complete stricture in the oral cavity (Grunwell & Harding, 1998), resulting in no audible oral release.

Nasal turbulence is defined as a 'snorting' or turbulent noise resulting from the approximation but inadequate closure of the superior border of the velum and the posterior pharyngeal wall (Duckworth, Allen, Hardcastle & Ball, 1990). This has been described as a 'nasal rustle' resulting from friction produced when an airstream passes through a small velopharyngeal gap (Kummer et al., 1992).

A *velopharyngeal fricative* is a 'snorting' or turbulent noise, resulting from the approximation but inadequate closure of the superior border of the velum and the posterior pharyngeal wall, and functions as a substitution for another sound (Duckworth et al., 1990). There is complete oral stricture during sound production with no audible oral release.

Nasal airflow errors have been described in varying ways by many authors (Peterson, 1975; Duckworth et al., 1990; Kummer et al., 1992; Sell et al., 1994). For the purpose of this study working definitions have been developed based on the definitions presented here. These working definitions and the discussion regarding controversy of definitions are presented in Chapter 4, Section 1.

Specific articulatory errors have been reported to be associated with cleft palate and/or velopharyngeal dysfunction (McWilliams et al., 1990; Sell et al., 1994; Wyatt et al., 1996). Sell et al. (1994; 1999) describe cleft type articulatory characteristics of speech including: dentalization; lateralization; palatalization; gliding of fricative/affricates; backing; weak articulation and nasalizaton, as well as compensatory errors. Dentalization or fronting of alveolar sounds have been noted in cleft speakers (Eurocleft Speech Group, 1993). Lateralization and palatalization have been reported to be the most common cleft type articulatory errors in children with cleft palate (Albery & Grunwell, 1993). Lateralization occurs usually on alveolar/velar targets, where the airstream is released centrally and laterally. Lateral realization, where there is total lateral release of a sound, can also occur. Palatalization occurs when there is secondary articulation of open approximation between the front of the tongue and the hard palate, simultaneously with a primary articulation of an alveolar sound. Palatal realization of sounds can also occur. According to Sell et al. (1999), gliding of fricatives and affricatives is a common characteristic but no research data associates gliding with compensatory articulation. Backing of sounds, alveolar targets backed to velar, or alveolar/velar to uvular sound place of articulation is a common characteristic of cleft speech. It is thought to be the product of abnormal neuromotor learning caused by structural abnormality (Wyatt et al., 1996). Weak articulation has been described as an articulation pattern associated with velopharyngeal dysfunction. This occurs when oral pressure is reduced but there is normal place of articulation (Wyatt et al., 1996). The reduction in pressure may be due to incomplete closure of the velopharyngeal sphincter and leakage of air through the nasal cavity. Perceptually, the pressure consonant may be soft and weak resulting in reduced intelligibility (Wyatt et al., 1996). Nasalized consonants can also be present when there is incomplete velopharyngeal closure. Nasalization occurs when the sound produced in the oral cavity is accompanied by nasal resonance (Grunwell & Harding, 1996). Nasalization normally occurs in English in the presence of nasal consonants; however, when produced on other sounds in the absence of nasal consonants, it is usually associated with velopharyngeal dysfunction.

Compensatory articulation errors are a distinct category of errors commonly found in patients with velopharyngeal dysfunction (Trost, 1981; Bzoch, 1989; Sell et al., 1994). Glottal and pharyngeal sounds are produced as compensatory articulations producing abnormal articulation at a place where success is more likely (Sell et al., 1994). Plosives, fricatives and affricates are produced below the inadequate velopharyngeal sphincter and therefore the velopharyngeal sphincter becomes irrelevant. Compensatory articulation includes: the glottal stop; pharyngeal stop; pharyngeal fricative and glottalization; The glottal plosive occur when there is closure of the vocal cords, with the build-up of pressure below the glottis; a pharyngeal stop is produced when the back of the tongue articulates with the posterior pharyngeal wall and there is a build-up of pressure which is released as a plosive. A pharyngeal fricative is produced at the same site of articulation; however, there is sufficient narrowing of the airstream to produce friction. The glottal stop can be used as a substitution for voiced and voiceless plosives and is usually used consistently across a variety of phonetic contexts (Sell et al., 1994). Glottalization or glottal coarticulation can also be present as compensatory articulation. The glottalized sound occurs when a plosive is produced with correct oral closure and secondary closure at the glottis. The oral closure is released, followed by release of the glottal closure (Kittelson, Broen & Moller, 1983). This articulatory pattern is also referred to as double articulation (Trost 1981; Wyatt et al., 1996) and glottal reinforcement (Sell et al., 1999).

In the assessment of speech problems associated with cleft palate, non cleft velopharyngeal dysfunction and nasal obstruction, analysis must be made at both segmental and suprasegmental levels, as problems at one or both levels distort speech and each has different implications for treatment. The focus of the present study is on the assessment at the suprasegmental level of nasality, specifically hypernasality and hyponasality, and the segmental nasal airflow errors in speech.

Children with cleft palate and/or velopharyngeal dysfunction are at risk for laryngeal/voice disorders, which may be congenital or behavioural (D'Antonio & Scherer, 1995). Abnormal voice patterns such as breathiness, abnormal pitch and laryngeal stridor are associated with syndromes such as Apert syndrome and Velo-Cardio-Facial syndrome (Witzel, 1995). Voice disorders may also be secondary to velopharyngeal dysfunction. D'Antonio, Muntz & Province (1988) reported that 41% of 85 patients with velopharyngeal dysfunction had abnormal voice characteristics, observable laryngeal abnormalities, or both. The Clinical Standards Advisory Group (CSAG) (1998) found that 29% of five year olds and 17% of twelve year olds assessed were dysphonic (Appendix 1).

Descriptions of vocal problems include hoarseness, breathiness, abnormal pitch, soft voice and strangled voice (McWilliams et al., 1990). These problems, according to McWilliams et al. (1990), may be a direct response to abnormal anatomy. Bzoch (1989) described the phonatory pattern associated with velopharyngeal dysfunction as weak and aspirate phonation. This weak phonation may be an attempt to minimise the resonance distortion from velopharyngeal dysfunction through weakening the fundamental frequency of the voice (Bzoch, 1989). Hoarseness may result from hyperfunctional use of the vocal cords and this may cause vocal cord oedema, nodules or other pathology (Witzel, 1995). The use of soft voice may be related to the inability to create sufficient oral pressure in speech, while a strangled voice quality may be associated with insufficient mouth opening during speech. Both of these abnormal vocal patterns may be unintentionally adopted to mask the presence of hypernasality or nasal emission (Witzel, 1995).

Acoustic and aerodynamic changes have been noted in individuals with velopharyngeal dysfunction and voice disorders (Appendix 1). Children with strained and strangled voice quality associated with velopharyngeal dysfunction were found to have high transglottal pressure and resistance, while children with breathy voice patterns were found to have low transglottal pressure (Lewis, Andreassen, Leeper & Macrae, 1993). Sapienza, Brown, Williams, Wharton and Turner (1996) reported that an individual was unable to increase intensity and had an increase in fundamental frequency when the velopharyngeal opening was greater than 20 mm. Zajac and Linville (1989) found significantly raised jitter (cycle-to-cycle variations in fundamental frequency) in children with velopharyngeal dysfunction. They also found a positive correlation between jitter and perceived nasality and a positive correlation between shimmer (cycle-to-cycle variations in amplitude) and hoarseness.

The complexity of speech problems asssociated with cleft palate and velopharyngeal dysfunction is emphasized in the literature. The importance of reliability, agreement and validity of speech assessment is well recognised (McWilliams et al., 1990; Cordes 1994; Wirz & Mackenzie Beck, 1995). In section 2.1, definitions of reliability, agreement and validity for general assessment will be reviewed and an outline of their use in previous research and in the present study will be provided. The terms, cut-off values, test sensitivity and test specificity will be defined within the context of recent research (section 2.2). The remaining sections of this chapter review three approaches to assessment of nasality and nasal airflow; perceptual, acoustic and aerodynamic. Each assessment technique will be considered and cut-off values, test sensitivity and specificity will be discussed.

2.1 Reliability, Agreement and Validity

The term reliability has been used with different meanings in the literature on the observational data in speech pathology. According to Cordes (1994), defining reliability requires differentiation between two distinctly different uses of the term. One use is a broad concept related to generally trustworthiness of obtained data. This definition refers to the dependability or reproducibility of test scores. The second use of the term refers to a mathematical relationship between an observed test score and the test takers 'true score' on the test. This means that observed scores are reliable if they vary with variations in true score rather than with variations in measurement error. Further confusion is caused by equating the definitions of reliability with estimates of reliability. The estimate of reliability refers to the manner in which reliability is evaluated. The most common methods of estimating reliability include:

- intra-rater reliability, which is often measured by calculating correlation or percent of agreement between measurements made by an individual rater;
- inter-rater reliability measures a percent agreement or correlation between different raters;
- test-retest reliability measures agreement of measurements from two different occasions (Streiner & Norman, 1995).

In the review of the literature on reliability studies, estimates of reliability for studies will be compared.

It is important to distinguish between the concepts of reliability and agreement. Reliability has been equated with inter-rater agreement, and this, according to Cordes (1994), has led to more confusion. Kreiman, Gerratt, Kempster and Berke (1993) state that agreement implies that two listeners assign identical meanings to each scale point, such that the concept of severe hypernasality has the same meaning for each listener as has the concept of mild hypernasality and normal nasality. In the literature, the percent of agreement between two assessments has been used to evaluate reliability of measures of speech problems associated with cleft palate (Albery & Grunwell, 1993; Sell et al., 1994; Wirz & Mackenzie Beck, 1995). Reliability implies that listeners rate voices in a parallel fashion without implying that each scale value has the same meaning. Correlation coefficients have been used as one way of reporting reliability (Watterson et al., 1993; Sell & Grunwell, 1993).

Distinction between the terms reliability and validity also needs to be made. Validity, according to Cordes (1994), refers to whether measurements actually measure what they are designed to measure. Validity, like reliability, refers to a broad concept which historically included content validity, criterion validity and construct validity. These three types of validity were defined by Streiner and Norman (1995). Content validity means that a scale or test contains sufficient items and adequately covers the domain under investigation. Criterion validity relates one measurement to another measurement of the same problem, usually a 'gold standard' which has been used in the field. Construct validity is more complex. This allows us to make inferences about a person based on a theory or construct. The theory explains the relationship between various behaviors. More recently a broader concept of validity has been used which refers to a process of hypothesis testing (Streiner & Norman, 1995). In general two questions need to be answered according to Streiner & Norman (1995):

- Does a test measure what it was designed to measure ?
- Can we make inferences with confidence about the people being tested?

In order to answer these two questions, test sensitivity and specificity are evaluated in the literature using cut-off values.

2.2 Cut-off Values

Cut-off values for test scores have been determined to distinguish normal and abnormal speakers. In perceptual studies, cut-off values can be a point on a scale where the listener decides that speech is abnormal. In some studies, the cut-off value on the scale was predetermined (Watterson et al., 1998a), whereas in other studies the cut-off value was decided post-hoc (Dalston, Neiman & Gonzalez-Landa, 1993).

For the Nasometer, cut-off values have been determined using scores from normal and clinical populations (van Doorn & Purcell 1998). In earlier studies (Fletcher, Adams & McCrutcheon, 1989), the cut-off values were determined by using two standard deviations above or below the mean for a normal population. Van Doorn and Purcell (1998) state that studies using clinical cut-off values (determined by comparing perceptual measures and nasalance scores) differ in their findings. Cut-off values may differ for different dialects and language as studies have shown that normal nasalance scores vary according to language and dialect (Seaver et al., 1991; Trindade et al., 1997). It would appear that using two standard deviations around the mean for a normal population and testing this on a clinical population may give a better indication of cut-off values or range of values. Van Doorn and Purcell (1998) point out that cut-off values should be used as a guide to the nasalance value that corresponds to the perception of normal or abnormal nasality.

Cut-off values for aerodynamic measurements were estimated using measurements from participants with repaired cleft palate, velopharyngeal dysfunction and model analogues of the vocal tract (Warren, 1967; Warren, 1979). In the model studies, a mechanical model of the vocal tract was developed which simulated velopharyngeal opening. Opening of the model velopharyngeal port ranged from 0 to 1.0 cm^2, thus representing adequate and inadequate closure of the port. Difference in pressure measurements between the nose and the mouth were compared to the size of the velopharyngeal opening in order to estimate cut-off values indicating velopharyngeal inadequacy.

It is evident, however, that using a cut-off value, whether it is determined from normal or clinical studies, has one major drawback. A small number of participants will be misclassified as normal or abnormal using a single cut-off value. Test sensitivity and specificity will provide an estimate of the number or percentage of cases misclassified and provide some validation measurement of the test.

2.3 Sensitivity and Specificity

Test sensitivity is the percentage of participants who are identified as having abnormal speech on one test and who are also identified as having abnormal speech on another test. Test sensitivity for the Nasometer is the percentage of participants judged to be hypernasal using a perceptual test and who are also identified as abnormal on the Nasometer (Watterson, Hinton & McFarlane, 1996). Test specificity refers to the percentage of patients who are identified as having normal speech on one test and who are also identified as having normal speech on another test. Using the Nasometer, the test specificity is the percentage of participants who are judged as having normal nasality on a perceptual scale and who also score within normal limits on the Nasometer (Watterson et al., 1996). An overall efficiency rating can be calculated by adding the number of times one test agreed with the other test, divided by the total number of opportunities (i.e. total number of participants tested) (Watterson et al., 1998a). In recent studies, a cut-off value has been determined for an instrumental measure, which distinguishes between normal and abnormal speech, and the results have been compared to other assessments of the same speech parameter. In this way, agreement between measurements can be ascertained.

Cut-off values, sensitivity and specificity will be discussed in relation to instrumental assessment of nasality and nasal airflow in a later section.

2.4 PERCEPTUAL, ACOUSTIC AND AERODYNAMIC APPROACHES

Three approaches to assessment of nasality and nasal airflow errors in speech are presented in this study. Perceptual assessment involves perceptual judgements regarding the presence or absence of nasality and nasal airflow problems, and severity of the problem. Acoustic assessment involves the measures of acoustic energy resulting from the coupling of the oral and nasal cavity (Moon, 1993). The most popular clinical and research acoustic technique is Nasometry, which measures oral and nasal acoustic signals. Aerodynamic assessment techniques have been developed to assess both nasality and nasal airflow errors. Various approaches are described below, with specific reference to the pressure/flow technique developed by Warren (1979). The pressure/flow technique has been described as a potentially powerful tool in the evaluation of velopharyngeal function (Moon, 1993).

2. 5 Perceptual Assessment

This section examines why perceptual judgements are fundamental to all other methods of speech assessment. It outlines the importance of reliability and validity, and the factors which influence reliability and validity. Scalar judgements of nasality and nasal airflow errors, and articulatory assessments of nasal airflow errors are discussed. Reports of reliability studies are evaluated and implications for future perceptual assessments are outlined.

Arguments have been made since the 1960's to support the use of perceptual judgements as a basis of assessment of voice and resonance. The central role of voice quality perception in the assessment of disordered speech is not surprising considering that the ultimate goal of speech is communication, and the ultimate test of acceptability of speech involves the perceptual acceptability to the listener (Gerratt, Kreiman, Antonanzas-Barroso & Berke, 1993). Kuehn (1982; p. 518) summarises the basis for perceptual judgements of speech when he states, that "in a sense, a speech disorder does not exist until it is perceived by a listener". Perceptual assessment is not only recognised as the primary approach to voice and resonance assessment, but has also been used to validate other assessment approaches. Auditory perceptual judgements of speech production of normal and disordered speech performance are central to the interpretation of all other forms of analysis (Folkins & Moon, 1990). According to Young (1969; p. 135) "a measurement of speech disorder is primarily a perceptual event and the observer's response necessarily represents the final validation for any measures". Therefore, if instrumental assessments are used to assess resonance, then perceptual judgements are used to validate the instrumental results. Perceptual judgement has been described as the only direct and logically valid measure of nasality (Moll, 1964). The perceptual judgement of nasality has also been described as the only indirect measure of velopharyngeal valving of critical importance (Shprintzen, 1995)[1]. It is now well recognised that the trained examiner's ear is indispensable in the assessment of nasality and that perceptual judgement of the Speech and Language Therapist is the first step in differential diagnosis of velopharyngeal dysfunction (LeBlanc & Shprintzen, 1996)

Despite its importance, perceptual assessment of speech has major disadvantages. Perceptual measures are difficult to calibrate and judgements may depend on experience and training (Kuehn, 1982). There may be variation in perceptual judgements between groups of listeners as well as between individual listeners (Kuehn, 1982; McWilliams et al., 1990). Phonetic factors and articulatory errors can influence the perception of nasality (Bzoch, 1989).

Kent (1996) expressed concern about auditory perceptual judgements in clinical practice because "the assumptions underlying the use of these judgements do not always characterise reality" (p. 7). He highlighted the following assumptions that are made regarding trained listeners:

- they have a common understanding of the meaning of perceptual labels;
- they use the same descriptors and scales to assess a speech sample;
- they can isolate one perceptual dimension from another;
- they have uniform reliability in judging dimensions of speech;

---

[1] The distinction, between direct and indirect assessment of velopharyngeal valving, was made by Shprintzen (1995). He argued that direct measures included nasendoscopy and videofluoroscopy, whereas indirect measures included nasometry, pressure/flow measures and auditory perceptual assessments.

- they make perceptual judgements for which inter-judge differences are smaller than differences needed for clinical classification.

But Kent (1996) argued that:

- listeners do not have equivalent definitions of dimensions to be rated;
- they fail to reach a consensus on what labels and scales should be used in clinical assessment;
- values that are obtained for one dimension may be influenced by other dimensions of speech that are inter-correlated;
- various perceptual ratings are not rated with uniform reliability; and
- differences among judges are larger than differences needed for diagnostic classification.

As a result of these problems the reliability of perceptual assessment of nasality has been questioned (McWilliams et al., 1990).

The difficulties of perceptual assessment are important factors affecting the reliability and validity of nasality and nasal airflow error assessments, and will be reviewed below.

2. 5. 1 Reliability, Agreement and Validity of Perceptual Assessments

Most reliability assessments of nasality and nasal airflow have been carried out using intra- and inter-rater reliability procedures (Watterson et al., 1993; Sell et al., 1994; Wirz & Mackenzie Beck, 1995). For example, Wirz and Mackenzie Beck (1995) defined the reliability of the Vocal Profile Analysis Scheme (which includes a section on nasality and nasal emission) as the ability to identify neutral or non-neutral parameters and to rate the degree of severity of that parameter to within one scalar degree of the 'right answer'. Intra-rater reliability of perceptual judgements is the ability of a listener to identify the parameter of speech as normal or abnormal and to rate the degree of abnormality in the same way during two different assessment sessions. Inter-rater reliability is the ability of two or more listeners to identify the same speech as normal or abnormal and to rate the degree of abnormality in the same way. In most studies agreement to within one point on a scale is considered to be acceptable.

Reporting agreement or rank correlations present some problems. Firstly, reporting agreement can produce artificially high reliability if there is a high occurrence of a clinical behaviour during an assessment of the reliability. For example, if a scale for assessment of nasality is used on a population which has a high proportion of participants with normal nasality, then the level of agreement of listeners will be high. This is because the judges may have found it easy to agree on the normal nasality, but may have found it more difficulty to agree on the abnormal nasality (Kearns & Simmons, 1988). As a result, the overall reliability coefficient may have been artificially inflated. Secondly, rank correlation coefficients only imply that the raters rated the speech samples in a similar manner. It does not take into account the exact agreement on the scales. Thirdly, the influence of chance agreement of ratings is not considered in either percentage of agreement or correlation coefficients. Listeners can agree on a rating by chance, and this needs to be taken into account when evaluating reliability. The use of occurrence reliability coefficients is suggested by Kearns and Simmons (1988) to overcome this problem. The occurrence reliability coefficient is calculated by including a formula for chance agreement in the estimate of reliability. This type of statistical agreement is used in kappa ratings of reliability. The kappa statistic relates the actual measure of agreement with the degree of agreement which would have been attained had the ratings been made at random (Bulman & Osborne, 1989).

In a review of the perceptual evaluation of voice quality[2], Kreiman et al. (1993) made two important recommendations based on the literature and clinical practice:

- both percentage of agreement and reliability should be reported;
- consideration should be made for chance agreements.

This may help overcome some of the problems with the use of the term reliability and with estimates of reliability.

---

[2] In recent literature, the discussion of perceptual assessments reviews all aspects of voice quality, including resonance and nasality.

Perceptual ratings of nasality have been reported to have high validity (D'Antonio & Scherer, 1995). However, other authors have questioned the validity of perceptual judgments of nasality (McWilliams et al., 1990; Kent, 1996). It is well recognized that speakers with velopharyngeal dysfunction can present with associated speech problems such as articulatory errors and phonation disorders. Listeners may have difficulty separating nasality from these other speech characteristics (Bzoch, 1989). It may be that raters rate their global perceptions and not the single element of nasality (McWilliams et al., 1990). Validity of perceptual assessments may be improved with training in the use of the assessment profile and comparing results to normal speakers (Sell et al., 1999).

Perceptual assessments of nasality have been used to assess validity of instrumental assessments of acoustic energy or air pressure/flow measurements (Haapanen, 1991a; Warren, 1967). According to Dalston and Warren (1986), it is generally agreed that perceptual judgments are the most appropriate standard against which to test the diagnostic accuracy of any objective diagnostic instrument, for example in studies of Nasometry and perceptual judgments (Paynter, Watterson & Boose, 1991; Nellis, Lehman & Lehman 1992). According to Kuehn (1982), instrumental techniques should be validated with perceptual measures as the 'common denominator', because, as previously described, this provides the only valid measure of nasality, since the listener is always an integral part of the communication process (Moll, 1964). However, using perceptual assessments of nasality and nasal airflow to validate instrumental assessment techniques is problematic. In this situation, perceptual assessment of nasality is assumed to be the 'gold standard' against which other assessment techniques can be compared. It has been shown that the perceptual assessment of nasality and nasal airflow can have poor reliability, weak agreement and poor validity due to the factors outlined in section 2.5.2. These problems will need to be overcome if perceptual judgments are to be used to validate instrumental techniques.

2. 5. 2 Factors Affecting Reliability/Agreement And Validity

There are several interacting factors which have been reported to affect reliability, agreement and validity, and these will be discussed below. Reliability in the studies reported in this section was evaluated using correlation procedures.

2. 5. 2 (i) Phonetic Aspects

Early studies of cleft palate speech indicated that high vowels are perceived as being more nasal than low vowels and front vowels are perceived as being more nasal than back vowels (Spriesterbach & Power, 1959; Lintz & Sherman, 1961; Carney & Sherman, 1971). Henningsson and Hutters (1997) also reported that high vowels are more vulnerable than low vowels to the velopharyngeal mechanism. This means that high vowels may be perceived as more hypernasal than low vowels in the presence of velopharyngeal dysfunction. Hence, the height of vowels in the speech sample may influence the perception of nasality.

Another early study reported that the phonetic context of a vowel influences perception of hypernasality. Lintz and Sherman (1961) found that in nasal and non nasal speakers:

- vowels in voiced environments were judged to be more severely nasal than vowels in voiceless environments, and
- vowels in fricative environments were judged to be more nasal than vowels in plosive environments.

More recently, McWilliams et al. (1990) stated that the phonetic context of the vowels may influence perception of nasality particularly in cases of borderline or marginal velopharyngeal dysfunction.

D'Antonio and Scherer (1995) reported that young children, who have a limited sound system consisting mainly of vowels and nasal consonants, may sound more nasal. This is probably due to the increased use of nasal consonants in speech.

The reliability of perception of nasality in isolated vowels has been investigated. Counihan and Cullinan (1970) reported a consistent trend for disparity of ratings to decrease and reliability coefficients to increase as the stimulus changed from isolated vowel to syllable to connected speech. Poor correlations for perception of nasality on isolated vowels with perception of nasality for the same vowel in connected speech have also been reported (Spriesterbach & Power, 1959; Carney & Sherman, 1971). In a more recent study, Bassich and Ludlow (1986) assessed the reliability of perceptual ratings of 13 dimensions of voice quality. One of the dimensions was nasality. They rated audio recordings of the vowels /a/ and /i/ on a 7 point scale. Participants included normal speakers and pathological speakers. They reported correlations between four judges' ratings of nasality for normal speakers (0.83) and for pathological speakers (0.63). One of the possible reasons for weaker correlations for the pathological speakers may have been that the nasality problems were not perceived on isolated vowels. Another possible cause was that the production of vowels in isolation may not have affected velopharyngeal valving to the same extent as the production of vowels in connected speech. McWilliams and Philips (1979) reported that velopharyngeal function may deteriorate as the complexity of the utterance increases in children with velopharyngeal dysfunction.

2. 5. 2 (ii) Vocal Quality

It is well recognised that children with repaired cleft palate and/or a history of velopharyngeal insufficiency may also present with voice/laryngeal changes. Varying compensatory laryngeal behaviours have been reported. Leder and Lerman (1985) suggested that these transglottal changes resulted in a reduction of nasal emission when velopharyngeal insufficiency was present. Studies indicate that laryngeal changes are required to maintain vocal pitch and loudness (Hamlet, 1973; Tarlow & Saxman, 1970) and that these changes may result in voice and laryngeal pathology (Sapienza et al., 1996). Such changes in voice quality may influence the perception of nasality and nasal emission. Changes in pitch have been shown to result in small changes in perception of nasality (McWilliams et al., 1990). LeBlanc and Shprintzen (1996) reported that hoarseness tends to mask hypernasality so that when velopharyngeal dysfunction and severe hoarseness occur, the hoarseness tends to be more noticeable than the hypernasality. Children with cleft palate and/or velopharyngeal dysfunction may use weak and breathy phonation, as this appears to minimise the distortion in resonance (Bzoch, 1989). Little is known about the exact effect of laryngeal/voice variations on the perception of nasality and nasal airflow errors and further research is required (Kuehn, 1982; McWilliams et al., 1990).

2. 5. 2 (iii) Articulatory Errors

Articulatory errors associated with cleft palate and velopharyngeal dysfunction have been described in detail in recent literature, with specific emphasis on errors of place and manner (Albery & Grunwell, 1993; Sell et al., 1994; Wyatt et al., 1996: Sell et al, 1999). Placement errors include: imprecise tongue tip movements; double articulation; lateral and palatal articulation; backing; and compensatory articulation such as glottal/pharyngeal realisations (Trost, 1981). Manner errors include: weak articulation; lateralization and palatalization; nasalization of consonants; and nasal airflow errors (Eurocleft Speech Group, 1993; Wyatt et al., 1996; Sell et al., 1999). It is well recognised in the literature that all these articulatory errors influence the perception of nasality. For example, some authors state that the perception of nasality increases in the presence of compensatory errors such as glottal realisations (Bzoch, 1989; McWilliams et al., 1990). This may be due to the close association of glottal realisations and severe velopharyngeal dysfunction (LeBlanc & Shprintzen, 1996). Kuehn (1982) believes that listeners evaluating speakers with articulatory and nasality problems may respond to one broad perceptual dimension rather than two separate and distinct dimensions.

2. 5. 2 (iv) Listening Conditions

Early studies hypothesised that audio tape recorded speech played backwards would eliminate or reduce the articulatory and phonetic influences on perception of nasality. Sherman (1954) compared ratings of speech stimuli played forward and ratings of speech stimuli played backwards and found that ratings of judgements obtained from forward playback were as reliable as ratings from backward playback. Counihan and Cullinan (1970) reported on results from their study and other studies that there was a consistent trend for reliability of perceptual judgements of nasality to increase as the stimulus changed from backward play to forward play. The current thinking on backward versus forward play of audio recordings of speech is that backward play has been found to distort some characteristics of the speech signal (McWilliams et al., 1990) and that backward play is not representative of normal conversational speech.

The effects of listening conditions on speech ratings were assessed in a clinical setting by Moller and Starr (1984). Trained listeners rated speech under three listening conditions:

- face to face;
- observing patient through a mirror and listening via a sound system;
- listening to a tape recorder.

They found that there was no significant difference between ratings of resonance under the three listening conditions. However, McWilliams and Philips (1979) stated that audio tape recorded speech may alter the perceptual judgements of speech components and does not provide the visual clues that are present in the live evaluation. The loss of visual information is particularly important in the assessment of nasal airflow errors from audio tape recordings, as the nasal/facial grimace that is often associated with nasal airflow errors cannot be seen. The quality of the acoustic signal is another important factor in perceptual judgements of nasality (Kent, 1996). In earlier studies, analogue audio tape recorders were used to record speech and these did not have the high acoustic quality of the more recent digital audio tapes. It is possible that digital audio tapes will overcome some of the problems of acoustic quality of audio recorded speech.

2. 5. 2 (v) Listener Training

Experienced judges have been used in studies of perceptual judgement of voice quality, to establish satisfactory levels of inter-rater agreement. Kent (1996) questions the degree to which experience alone can be assumed to guarantee satisfactory inter-rater agreement. Kreiman et al. (1990) found that experienced and naive raters attended to different aspects of voice quality when judging voice. Naive raters all used similar perceptual strategies, while experienced raters used varying strategies which they considered to be important. The authors believed that the different strategies depended on their clinical experience. Experience is valuable if it ensures fundamental commonalities (Kent, 1996).

A 13-dimension perceptual rating system was used to train new clinicians to assess voice quality (Bassich & Ludlow, 1986). Despite extensive training they found that reliability of inexperienced judges was lower than that of experienced judges. They concluded that reliability may be increased by training judges intensively on dimensions of voice to be assessed and by using anchor stimuli.

2. 5. 2 (vi) Group Judgements

Reliability of nasality ratings has been found to improve considerably when groups of listeners rated speech samples and the mean scale values were used to calculate agreement (Counihan & Cullinan, 1970). McWilliams et al. (1990) state that scales are reliable when groups of listeners participate and mean group ratings are used in the final ratings. Group rating is not practical however, in the clinical situation. Group ratings may be used to assess intra-rater reliability on a regular basis. For example, McWilliams et al. (1990) suggested the following approach: the listener tests his/her reliability against group ratings by listening to and rating speech samples previously rated by a group of listeners. The listener then discusses the ratings in comparison with the group rating and attempts to agree with group ratings. If listeners can agree with each other and frequently re-establish their agreement, they may have confidence in their ratings.

Recently, Watterson, Lewis and Dalston (1998b) stated that the concept that data from a panel of raters is more reliable than data from a single rater pair is a misconception. This misconception is based on misleading statistics. It was pointed out that using variance based statistics, such as intra-class correlations, previous researchers reported good agreements. But, a function of intra-class correlations is that the correlation coefficient increases as the number of raters increase. By comparing intra-class correlations and kappa statistics, Watterson et al. (1998b) indicated that a good intra-class agreement of 0.80 for 5 raters had a weak multi-rater kappa score of .03. They also highlighted the fact that correlations do not measure agreements. They concluded that group ratings are no more reliable than individual ratings.

2. 5. 2 (vii) Summary and Implications

The literature indicates that a speech sample used in assessment should include sounds that are vulnerable to increased nasality (Section 2. 5. 7 Speech Stimuli). The vowel height is an important aspect in the perception of nasality and this may be used to identify various degrees of hypernasality. In some cases nasality may be perceived on high vowels and not on low vowels. This distinction should be considered both in the speech stimuli used for assessment and in the perceptual analysis of speech. A description of voice quality should be noted during perceptual analysis of nasality, and should include the presence of dysphonia, abnormal pitch and reduced volume. This information on voice quality, plus a detailed phonetic/phonological analysis, may help the listener to separate these speech parameters from nasality and nasal airflow errors in order to complete a reliable and valid assessment. Training listeners and evaluation of listening conditions (comparing live analysis with analysis using high quality audio tape recordings of speech) may also help improve reliability and validity.

Psychological scaling techniques have been developed in an effort to reduce some of the problems associated with perceptual assessment of speech. However, many of the factors that influence reliability and validity also influence scalar judgements.

2. 5. 3 Scaling Techniques

Three common scaling procedures have been used for the assessment of speech disorders: direct magnitude estimation (DME); equal interval scaling (EAI); and paired comparisons (Kuehn, 1982). With the direct magnitude estimation technique, listeners are presented with a speech sample usually representing the middle of the range of severity of a particular speech dimension. The standard sample is assigned a number and listeners are instructed to assign a number to subsequent samples in relation to the standard sample (McWilliams et al., 1990; Kuehn, 1982). Using equal appearing interval scaling, the listener is asked to assign a number for a speech sample on a 5, 7, or 9 point scale, where 1 usually indicates a mild degree of deviation from the norm and the upper most number on the scale indicates severe deviation. For paired comparison, the listener compares all stimuli in sets of two. Paired comparison, according to Folkins and Moon (1990), can be laborious and is not easily applicable to speech judgements. They point out that although DME is sometimes more reliable than EAI scaling, the differences between the two procedures may not always justify the amount of work required for DME procedures in the clinic. Traditionally EAI scales are most frequently used in the assessment of hypernasality (Henningsson & Hutters, 1997)

Schiavetti (1992) outlines two major advantages of the use of scaling procedures in assessment of communication disorders. Firstly, the scaling of many dimensions of disordered communication is considered to be the most direct assessment of a particular dimension. Secondly, the scaling procedure is a simple technique which is available to a wide range of users in the clinical setting because it usually only requires a paper and pencil format rather than expensive equipment.

Some problems in the use of scales are evident. Firstly, on many scales the individual points are not clearly defined. For example, on Wilson's Buffalo Resonance Profile (1979), hypernasality, hyponasality and nasal emission are rated on a 7 point equidistant scale. But, Wilson does not attempt to define each of the scalar points on the scale. This means that replicability and inter-judge reliability cannot be guaranteed (Wirz & Mackenzie Beck, 1995). Secondly, various types of rating scales are available for rating nasality. Scales differ according to the number of points on the scale (McWilliams et al., 1990). As a result, different numerical values represent mild, moderate and severe for each type of scale. Thirdly, on some scales the lowest level on the scale is 'acceptable hypernasality whereas on other scales the lowest level is 'absent hypernasality' (Henningsson & Hutters, 1997). These differences cause confusion if nasality is compared using different scales. It also makes intercenter comparison difficult. These problems can result in poor reliability of scales. Another problem is that reliability increases as the number of points on the scale decrease (Henningsson & Hutters, 1997). McWilliams et al. (1990) reported that the more choices the listeners have to make, the harder it is to develop reliability. However, if the amount of detail of the scale is reduced, then relevant information may be lost (Henningsson & Hutters, 1997). Finally, the validity of rating scales for nasality can be a matter for concern if the listeners have difficulty separating hypernasality from other speech errors (McWilliams et al., 1990).

Some of the problems of scalar techniques may be overcome by defining each point on the scale clearly. It would be useful to consider a descriptive scale rather than a numeric scale. In this way, rather than assign a number to a speech sample, the listener assigns a description of the parameter being assessed.

2. 5. 4 Use of Anchor Stimuli

An important aspect of the perceptual assessment technique is the standard upon which judgements are based. Without a pre-established standard or suitable training, judgements of resonance disorders are likely to be unreliable (Kuehn, 1982). Anchor stimuli have been used to establish standards for perceptual assessment. This procedure involves presentation of auditory stimuli to the listener prior to the speech stimulus to be judged. The stimulus may be normal or a pre-established degree of abnormality. The listener then compares the stimulus against a pre-established standard. Anchor stimuli or reference samples have been used in different ways in the literature. Gerratt et al. (1993) used anchor stimuli that spanned the range of the voice quality being assessed (roughness) and were perceptually equidistant. Raters' judgements on an EAI scale were compared with and without anchor stimuli. When using the anchor stimuli, participants were asked to select the anchor stimulus that best matched stimulus for roughness. Raters were allowed to listen to the anchor stimuli and speech stimuli as often as was necessary before responding. Gerratt et al. (1993) found that listeners agreed significantly more frequently when anchor stimuli were used. However, it is argued that using EAI scales in this way is more representative of a DME technique. The DME technique has been reported as being more reliable than the EAI technique (Folkins & Moon, 1990). It is possible that the use of a DME technique and not solely the use of anchor stimuli improves reliability.

Anchor stimuli have been used for training raters in the perceptual assessment of nasality and nasal emission. Karnell et al. (1985) presented the raters with two anchor stimuli. One represented mild hypernasality and/or nasal emission; the other represented severe hypernasality and/or nasal emission. The raters then rated 10 practice samples to familiarise themselves with the rating procedure. No feedback was given to the raters following the training ratings. Spearman Rank correlations for intra- and inter-rater reliability were calculated. Karnell et al. (1985) found a correlation coefficient of 0.82 for intra-rater reliability and 0.84 for inter-rater reliability. Unfortunately they did not report on the percent of intra- and inter-rater agreements or account for chance agreements. The rank correlation merely indicates that raters rated nasality and/or nasal airflow errors in a similar manner. Also, in this study nasality and nasal airflow errors were not separated as two distinct speech parameters and therefore the validity of the assessment procedure is questionable.

Experienced clinicians who frequently assess a specific type of disorder may develop their own standard references, which are stored in long term memory (Kuehn, 1982). Hence, a specialist Speech and Language Therapist may have built up a repertoire of references for different degrees of nasality and nasal airflow errors from previous clinical experience. It is because of this internal standard that experienced judges have been used for reliability ratings in many studies. In theory, the use of anchor stimuli should improve reliability of perceptual judgements. However, further research into the approaches to using anchor stimuli and the effect of anchor stimuli on reliability of perceptual judgements needs to be carried out.

2. 5. 5 Perceptual Assessments of Nasal Airflow Errors

Scalar judgements have been used primarily to assess nasality disorders, and have only occasionally been used for assessing nasal airflow errors (Wilson, 1987). The problems of scalar assessments outlined above also apply to scalar assessments of nasal airflow errors.

Much of the literature on perceptual assessment of speech and voice relates to assessment of voice characteristics of roughness, loudness, pitch and resonance (Kearns & Simmons, 1988; Bassich & Ludlow, 1986; Kreiman et al., 1993). Little has been written on perceptual aspects of judgements of nasal airflow problems. Bzoch (1989) raises the question of whether nasal emission should be considered as an articulation or a voice disorder. The problem of trying to fit the phenomenon of nasal emission into the conceptual framework of voice-versus-articulation disorders of speech has led to the use of various descriptive terms (Chapter 4. Section 2 Definitions of Terms). Bzoch (1989) goes on to suggest that recognition of separate nasal emission distortions of pressure consonants, and the accompanying distortion of vowel sounds as a resonance disorder, is critical in the differential diagnosis of velopharyngeal insufficiency. He suggests that nasal emission be considered as a separate categorical aspect of articulation.

In the 1960's and 1970's, articulation tests were developed specifically to assess the production of consonants most likely to be affected by velopharyngeal dysfunction (i.e., plosives, fricatives and affricates). These tests included the Iowa Pressure Articulation Test (IPAT) (Morris, Spriesterbach & Shelton, 1961) and the Error Pattern Articulation Test (Bzoch, 1989). These tests allowed for recording of error types which are important in differentially diagnosing speech problems associated with velopharyngeal dysfunction and problems associated with other aetiologies (McWilliams et al., 1990). However, the above tests only analysed words in isolation and not connected speech. As McWilliams et al. (1990) point out, children may achieve normal sound production at word level but not in conversational speech. For example, children with borderline velopharyngeal function, as described by Warren (1979), can obtain adequate velopharyngeal function in single words but inadequate velopharyngeal function during conversational speech, where rapid movement of the soft palate and pharyngeal walls are required.

Nasal airflow errors were described in an auditory-articulatory framework to aid in the phonetic/phonological analysis of cleft speech errors (Eurocleft Speech Group, 1993; Grunwell & Harding, 1996). In the Eurocleft study (1993), nasal airflow was one of five clusters of speech errors which were typical of cleft palate speech. Grunwell and Harding (1996) classified nasal airflow problems as 'cleft type articulatory' error, including audible nasal escape, nasal fricatives and nasal turbulence.

Nasal emission and nasal turbulence have been described as errors of manner of articulation, whereas nasal fricatives have been described as placement errors (Wyatt et al., 1996; Sell et al., 1999). In the Great Ormond Street Speech Assessment (GOS.SP.ASS.) (Sell et al., 1999), these descriptors of nasal airflow errors were used as part of an assessment profile. They allow for the description of nasal airflow under the articulatory error section as well as under a separate section for resonance and nasal airflow. The listener records the nasal airflow error in terms of place and manner, as a 'cleft type error'. The listener also makes an overall assessment of nasal emission and nasal turbulence by rating each on a 3 point scale. Nasal emission and nasal turbulence is recorded as audible/inaudible, consistent/inconsistent and accompanying/replacing consonants.

In 1998, Harding and Grunwell distinguished between active and passive cleft-type speech characteristics. Active processes were defined as "alternative articulations thought to have been actively generated in order to establish the necessary phoneme distinctions between individuals consonant targets" (Harding & Grunwell, 1998. p. 334). In other words, active processes established meaningful contrasts in speech. Passive processes were usually a result of structural defects or velopharyngeal dysfunction. During the production of passive realisations there were no alterations of the articulation patterns for the intended consonant. Using the active/passive conceptual framework of cleft type speech analysis, nasal emission was described as a passive cleft type error which accompanied sounds. Nasal fricatives were described as active or passive. Active nasal fricatives occurred when air was actively directed nasally as an alternative to an oral fricative. The airflow through the mouth was stopped by the lips or the tongue and directed through the nose. A passive nasal fricative was defined as an unreleased (s) double articulated with a lowered voiceless nasal. This was described as an intended /s/ with unintended nasal airflow. A passive nasal fricative could be converted into an oral fricative by holding the nose and preventing nasal airflow.[3]

These studies represent a major advancement in the analysis of perceptual nasal airflow problems associated with velopharyngeal dysfunction. They distinguished between various types of nasal airflow problems and provided a framework for a more detailed descriptive assessment of nasal airflow problems. The conceptual framework provided in these studies has overcome the problem identified by Bzoch (1989) of trying to fit the phenomenon of nasal emission into a conceptual framework of voice-versus-articulation disorders.

Some problems still exist in the assessment of nasal airflow problems. D'Antonio and Scherer (1995) stressed that assessment of nasal emission should include information regarding severity and frequency. In the GOS.SP.ASS, nasal emission and nasal turbulence are each rated on a 3 point scale, where 0 indicates absent nasal emission/nasal turbulence, 1 indicates slight emission/turbulence and 3 indicates marked or severe emission/turbulence. However, slight or severe may be influenced by consistency and frequency. There is no format in GOS.SP.ASS (Sell et al., 1994; 1999) for assessing frequency of nasal emission or nasal turbulence. The frequency is reflected in consistency where Sell et al. (1999) stated that if emission occurred occasionally or intermittently, it would be considered inconsistent. They do not allow for the situation where the level or frequency of nasal emission or nasal turbulence varies from one speech situation to another. Also, as this is a screening assessment, only two degrees of severity

[3]Occlusion of the nares has been used for assessment of hypernasality in the literature. Bzoch (1989) described his 'cul-de-sac' test for hypernasality. The tester asked the child to repeat each word twice. During the second production, the tester pinches the nares closed and notes a shift in quality from hypernasal to a cul-de-sac type nasality.

of nasal emission/turbulence are recorded. Further detailed assessment of nasal emission/turbulence may be required, which includes information on strength and frequency of the nasal airflow error. Although Sell et al., (1999) describe the use of a nasal fricative under cleft type articulatory errors they do not include nasal fricatives in the overall resonance/ nasal airflow section. The production of nasal fricatives may influence overall judgements of nasality and nasal airflow errors and this group of articulatory errors should be included in the general judgement of nasal airflow errors. Like nasal emission and nasal turbulence, a nasal fricative can be seen as an airflow error as well as an articulatory error.

2. 5. 6 Studies Reporting Estimates of Reliability of Perceptual Assessments

Varying estimates of reliability/agreement results have been reported in the literature. Using the Nasal Emission Test (Bzoch, 1989), nasal airflow was detected auditorily and checked using an airflow paddle held under the nose. The child was asked to repeat ten bisyllabic words containing the pressure consonant /p/ and evidence of airflow was recorded. Bzoch (1989) reports 96% agreement between independent listeners' rating of the presence or absence of nasal emission.

The agreement of judgements of articulation (including nasal emission) was evaluated by Philips and Bzoch (1969). They found good intra and inter-judge agreement between the presence and absence of articulation errors, but poor agreement between judges on the types of errors which occurred. In the nasal emission category inter-judge agreement ranged from 0% to 27%. This agreement is considerably lower than the above agreement. The authors point out that low agreement may be due to poor definition of criteria for classification of errors (See Chapter 4 Section 1, Definitions of Terms).

The percent of agreement between two raters on rating the presence/absence of nasality and nasal emission was reported by Sell et al. (1994). They reported 84% agreement for hypernasality, 92% agreement for hyponasality and 77% agreement for nasal airflow errors. Similar results were reported for agreement on ratings of nasality using 5 point scales. Lohmander-Agerskov (1996) found 70-80% agreement on ratings of hypernasality and nasal emission, while Watterson et al. (1998a) found 80% agreement on ratings of hypernasality.

Correlation coefficients have been used to report reliability of perceptual judgements of nasality. Intra-rater correlations have been reported using various correlation coefficients. A Spearman Rank correlation of 0.74 for intra-rater reliability of judgements of nasality was reported by Sell and Grunwell (1990). The Pearson product moment correlation coefficient has been calculated in several studies to evaluate intra-rater reliability. Hardin, Van Demark, Morris and Payne (1992) reported correlations ranging from 0.80 to 0.93 for ratings of hypernasality, and 0.9 to 1 for ratings of hyponasality using a 5 point scale. Watterson et al. (1993) reported correlations ranging from 0.76 to 0.88 for ratings of hypernasality. Using a 7 point equidistant scale, Ramig (1982) reported intra-rater correlation coefficients ranging from 0.55 to 0.99. Correlations ranged from 0.79 to 0.99 for ratings of hypernasality and from 0.66 to 0.88 for ratings of hyponasality using a 6 point scale (Pinborough-Zimmerman, Canady, Yamashiro & Morales, 1998).

Inter-rater correlations have also been reported by Sell and Grunwell (1990) who found a Spearman Rank correlation of 0.76. Varying Pearson correlation coefficients for ratings of hypernasality have been reported: 0.77 (Pinborough-Zimmerman et al., 1998); 0.82 to 0.92 (Hardin et al., 1992); and 0.96 (Watterson et al., 1993). Pearson correlation coefficients of 0.80 to 0.93 for ratings of hyponasality were reported by Hardin et al. (1992). Pinborough-Zimmerman et al., (1998) also reported an inter-rater correlation of 0.76 for ratings of nasal emission.

Good agreement was found in a study on inter-rater agreement of velopharyngeal settings using the Vocal Profile Analysis Scheme (Laver, Wirz, Mackenzie Beck & Hiller, 1981) (Sweeney, 1984). Raters rated judgements of nasality on a 6 point scale where ratings of 1 to 3 indicated normal nasality and ratings of 4 to 6 indicated hypernasality. Agreements between 3 raters to within 1 scalar point was found to range from 77% to 87%. However, a good agreement on this type of scale was not surprising as only 3 scalar points of hypernasality were used and agreement to within 1 scalar point was reported. Further analysis of rater agreement was carried out using Pearson correlation of exact agreement of ratings. This indicated weak correlations ranging from 0.02 to 0.35. Hence, although good agreement was found, detailed analysis did not report good correlations. Differences in agreement/reliability resulting from different statistical analyses are underscored in this study.

Most of the studies reported good agreement in ratings of nasality or good correlations between ratings. However, these results do not always indicate good reliability. The percent of agreement does not consider the level of chance agreement that may have occurred. Watterson et al. (1998b) reported that there was a 20% probability of chance agreement of ratings using a 5 point scale. This chance agreement was not considered in the above studies. They analysed data comparing ratings of two raters using a 7 point scale. Exact agreement between ratings was 20%; agreement within 1 scale point was 30%. They also found a Pearson correlation coefficient of 0.81. However, by calculating a weighted Kappa score to estimate agreement, taking account of chance agreement, the result was 0.36. Watterson et al. (1998b) highlighted the fact that reporting strong correlations can only indicate that raters rated nasality in a similar way. This shows that a strong correlation does not indicate good reliability.

Reliability studies, using the kappa analysis, were carried out by the Clinical Standards Advisory Group, Cleft Lip and/or Palate (1998). These studies used a modification of the Cleft Audit Protocol for Speech (CAPS) (Harding et al., 1997). The Cleft Audit protocol is a reduced version of the GOS.SP.ASS (Sell et al., 1994), which was developed for audit purposes. Raters underwent three days of intensive training in the use of the CAPS. Inter-rater reliability was excellent for ratings of hypernasality (kappa score of 0.8). Inter-rater reliability was weaker for hyponasality (kappa score of 0.57) and nasal emission (kappa score of 0.59). These results indicate that using a short, simple rating scale, there is moderate to excellent inter-rater reliability for assessment of nasality and nasal emission for trained listeners.

Results of reliability studies are conflicting and indicate considerable variation in intra- and inter-rater reliability scores for different studies. This may be explained by the use of different methodologies, scales and statistics used to estimate reliability in all the studies. In some of the above studies only the percent of agreement between raters was reported and no statistical reliability scores were calculated. Another problem was that the studies only evaluated the agreement for the presence or absence of nasality or nasal airflow errors and no results on agreement regarding the severity of the disorder was presented. Sweeney's (1984) study and the report by Watterson et al. (1998b) highlighted the variation in results when different approaches to assessment of reliability were used.

Problems of reliability testing and factors affecting reliability indicate the importance of the speech stimuli used in perceptual assessment of nasality and nasal airflow errors.

2. 5. 7 Speech Stimuli

As indicated above, there are problems with the assessment of nasality in isolated vowels and the assessment of nasal airflow errors in words in isolation. Considerable attention must be paid to the speech sample used in perceptual assessment. Many clinicians use a standard list of words and sentences to assess articulation, resonance and nasal emission (D'Antonio & Scherer, 1995). The phonetic makeup of these utterances is of utmost importance. Henningsson and Hutters (1997) stated that word structure, vowel height, stress, and phonetic context of the vowels, should be considered when selecting the speech stimuli. Sentences such as the Pittsburgh Sentences (Phillips, 1986) or the GOS.SP.ASS sentences (Sell et al., 1999) contain utterances which have varying proportions of nasal consonants to assess hyponasality and hypernasality (Appendix 2). They also contain high pressure consonants in different phonemic structures which are vulnerable to velopharyngeal dysfunction. This allows for detection of nasal airflow errors, weak intra oral pressure and compensatory articulation. Another consideration is that the sentences should be imageable and meaningful (Sell & Grunwell, 2000). Van Demark (1964) found a high correlation between sentence repetition and spontaneous speech. Sentence repetition is therefore considered to be an effective way of collecting a speech sample.

Consideration of the phonetic makeup of utterances is now accepted in the literature, and the importance of assessment at different levels of speech (i.e. word, sentences and conversational speech) has been stressed (LeBlanc & Shprintzen, 1996). Balanced representative data samples are preferable to gain a comprehensive assessment of an individual's speech (Grunwell, Sell & Harding, 1993). So, although it is important to elicit speech samples that are most vulnerable to velopharyngeal dysfunction, it is also important to elicit a speech sample that is representative of conversational speech. This also provides the opportunity for the listener to evaluate the consistency of speech patterns in different speech situations.

2. 5. 8 Summary and Implications

The foregoing section has provided an overview of the importance of perceptual assessment of speech. (Young, 1969; Moll, 1964; Gerratt et al., 1993). However, significant problems have been found in the perceptual analysis of nasality and nasal airflow errors. Although scalar judgements of nasality and articulatory assessment of nasal airflow errors are used clinically and in research, many factors have been found to influence the perceptual judgements. Phonetic context, articulatory errors, variations in pitch and loudness, listener training and conditions have been found to affect reliability and validity of assessment procedures. By knowing the nature of the limitations of perceptual tests resulting from these influencing factors, the undesired effects can be minimised (Kent, 1996). It has been proposed that the use of anchor stimuli (Kuehn 1982; Gerratt et al., 1993), improved training in the specific dimensions of nasality and nasal airflow errors (Bassich and Ludlow, 1986) and supplementing perceptual judgements with instrumental analyses (Kent, 1996) may help to improve reliability, agreement and validity. These three factors and the problems identified in this chapter have been considered in the development of a perceptual profile for analysis of nasality and nasal airflow errors in disordered speech. (Chapter 4. Perceptual Framework). Specific recommendations identified from the literature include the need to:

- define the terminology;
- use speech samples with specific phonetic makeup at various levels of speech;
- assess nasality and nasal airflow errors in an articulatory and descriptive framework;
- report agreement and statistical reliability of perceptual judgements;
- evaluate the use of anchor stimuli; and
- validate perceptual assessments using instrumentation.

Instrumental techniques for the assessment of nasality and nasal airflow have been developed and improved during the past decade. The following sections review the literature on nasometric assessment of nasality and aerodynamic assessments of nasal airflow.

2. 6 Nasometry

This section describes the Nasometer, a computer-based instrument which provides the examiner with a quantitative value that represents relative amounts of oral and nasal acoustic energy during speech. The development of the instrument is briefly outlined. Research on the use of the Nasometer is reviewed under the following headings: speech stimuli; studies of normal speakers; clinical studies - reliability and validity using correlations between scores from the Nasometer and perceptual judgements of nasal resonance; test sensitivity and specificity and nasal obstruction. Conclusions and the need for further investigations are presented.

2. 6. 1 Description

The Nasometer was introduced in 1986 by Kay Elemetrics. It was designed to measure oral and nasal acoustic sound signals and calculate a score which represents the ratio of the energy in the two signals (Fletcher, Adams & McCrutcheon, 1989). "This microcomputer-based instrument employs microphones on either side of a sound separator plate which rests on the upper lip. The signal from each microphone is filtered and digitized by custom built electronic modules. The data is then processed by a computer and accompanying software. A numeric ratio of nasal acoustic energy to the sum of nasal plus oral acoustic energy is calculated, multiplied by 100 and expressed as a 'nasalance' score" (Dalston & Seaver, 1992; p. 17) (Figure 2.1).

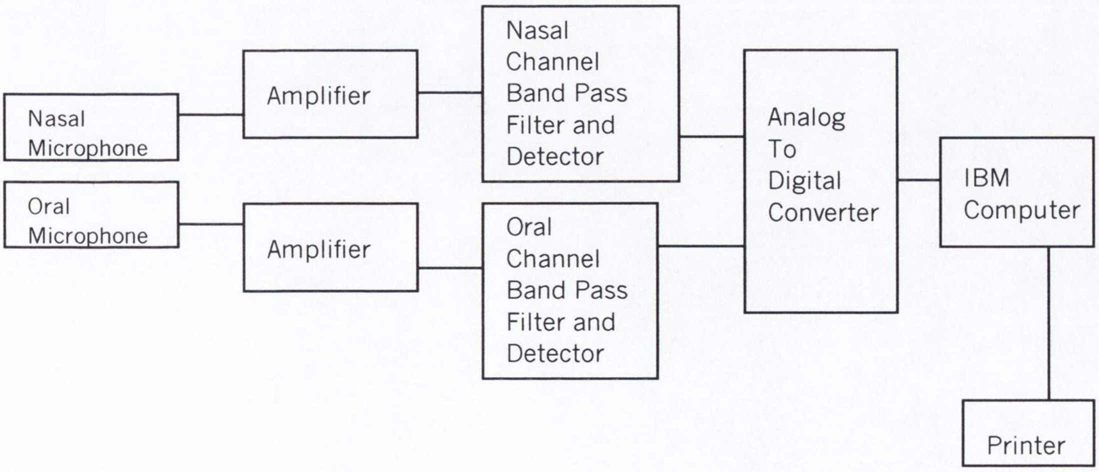


Figure 2.1. Block diagram of the Nasometer instrumentation (Fletcher et al., 1989; p. 249).

LaBlanc, Steckol and Cooper (1991) describe the Nasometer as a non invasive instrument, consisting of a headset, microprocessor, a printed circuit board and software. Fletcher et al. (1989) describe the software, which includes data acquisition, data editing and analysis, stimulus presentation, display generation, file management and various utilities. The statistical package in the data analysis and editing provides mean and standard deviation of nasalance, minimum and maximum nasalance values, time range and cursor values.

2. 6. 2 Background

The original version of the Nasometer, Tonar (The Oral Nasal Acoustic Ratio) was developed by Fletcher (1970). This machine contained a lead chamber which separated the mouth from the nose. Microphones were suspended in each chamber so that separate acoustic signals could be recorded. Signals were amplified and analysed using spectral analysis. The outputs were further analysed and displayed using a galvanometer/ oscillograph. Analogue computation of the oral/nasal ratio was also performed. Fletcher and Bishop (1970) reported a rank order correlation of 0.74 between measures of oral-nasal ratio measured by Tonar and ratings of hypernasality in a cleft palate population of 20 children.

The revised Tonar 11 was described by Fletcher (1976), who reported improved correlations. In Tonar 11 the following definition of nasalance was introduced:

$$\text{Nasalance} = \frac{N}{N+O} \times 100$$

(Fletcher et al., 1989)

This provides a nasalance score which may be compared to perceptual judgements of nasal resonance. The definition has several advantages, according to Fletcher et al. (1989). "The comparison between nasal output and the combined nasal-plus-oral outputs seems intuititively close to how listeners likely process the nasalized speech signals as they discern the degree of nasality" (Fletcher et al., 1989; p. 247). The other advantages are that the formula establishes limits at each end of the nasalance continuum and that the percentage scores are meaningful to the tester and the patient (Fletcher et al., 1989). Tonar 11 differed from Tonar in the use of frequency bandwidth of the oral and nasal signals. Fletcher (1976) found that the closest agreement between nasalance scores and perceptual judgements of nasality could be achieved when the band filter was centred in the region of 500 Hz, with a 300 Hz bandwidth around that central frequency. In this study, Fletcher compared perceptual judgements of nasality with nasalance scores. He found that perceptual judgements were highly variable and that highest agreement between the acoustic and perceptual measures was found when listener scores were pooled. The correlation between mean perceptual judgements of nasality and nasalance was 0.91.

The Nasometer is similar to the Tonar 11, but differs in structure, function and practical features. It contains a sound separator that separates the acoustic outputs from the nose and the mouth, electronic circuits for processing acoustic signals and transmitting them to a computer and a personal computer for receiving data, processing the information, calculating and displaying scores (Figure 2.1). A review of the speech samples used in nasometry will be presented in the section below.

2. 6. 3 Speech Stimuli

Three standard passages are provided in the manufacturer's manual for use with the Nasometer. These include: the *Nasal Sentences* Passage, which is loaded with nasal consonants; the *Rainbow Passage* which contains a mixture of oral and nasal consonants in the proportion found in everyday speech (Fletcher et al., 1989); and the *Zoo Passage,* which is devoid of nasal consonants (Appendix 3).

Dalston and Seaver (1992) assessed the relative value of the three standardized passages in the nasometric assessment of patients with velopharyngeal inadequacy. A series of 155 patients with various clinical diagnoses of cleft palate and craniofacial anomalies were studied. Nasometric scores obtained using the three reading passages were compared to perceptual judgements of hypernasality and hyponasality. Results indicated that the correlation between nasalance scores on the *Rainbow Passage* (phonemically balanced passage) and perceptual judgements of nasality was modest. Nasalance scores on the *Nasal Sentences* (nasally loaded passage) were significantly related to perceptual judgements of hyponasality. Nasalance scores on the *Zoo Passage* (non nasal passage) were significantly related to perceptual judgements of hypernasality. These results confirmed previous studies which indicated that nasalance scores on the *Zoo Passage* were useful in identifying patients with velopharyngeal dysfunction (Dalston Warren & Dalston, 1991b), while scores on the *Nasal Sentences* aided identification of participants with hyponasality (Dalston et al., 1991a). The authors concluded that the *Rainbow Passage* was of no additional value in providing clinical information regarding velopharyngeal dysfunction or upper airway patency. In this study, perceptual judgements of nasality were made on conversational speech, whereas Nasometry was based on the *Rainbow Passage*. As the *Rainbow Passage* was a phonemically balanced passage, one would have expected it to correlate well with conversational speech, but this was not the case. Dalston and Seaver (1992) suggested that as the *Rainbow Passage* was syntactically more complex than the other passages, it may have been more difficult for children to read or repeat, thus affecting scores on the Nasometer. A drawback of this study is that only one listener judged resonance and no intra- or inter-judge reliability study was carried out.

Watterson et al. (1996) developed two speech samples which were designed to parallel the *Zoo Passage* (devoid of nasal consonants) and the *Rainbow Passage* (phonemically balanced, with 11% nasal consonants). The former passage was called the *Turtle Passage* and the latter was called the *Mouse Passage* (Appendix 4). These two passages were syntactically and semantically easier than the original passages, as Watterson et al. (1993) had previously found the original passages too difficult for young children to recite. There was no significant difference between mean nasalance scores for normal children during the production of the *Turtle Passage* and during production of the *Zoo Passage*, thus indicating no difference in scores when simplified passages were used. Unfortunately, mean nasalance scores on the *Mouse Passage* were not compared to mean nasalance scores on the *Rainbow Passage*. However, correlations between nasality ratings and nasalance scores on the *Mouse Passage* were low, as were correlations between the *Rainbow Passage* and nasality ratings in the Dalston and Seaver study (1992).

A simplified Nasometric assessment was devised by MacKay and Kummer (1994) for use with children. This was provided with the manual of the upgraded Nasometer (version 6200.3). The speech sample in this test included syllables and sentences which were to be repeated by the children, and two short passages which were read. One passage contained approximately 8% nasal consonants, while the second passage was devoid of nasal consonants (Appendix 5). Norms for each speech sample were provided in the manual. One major problem with MacKay and Kummer's speech samples was that the reading passages were found to be culturally biased and therefore difficult for non-American children to read or repeat. (Normal Nasometry Pilot Study, Appendix 6). An example of a culturally biased passage is the *Bobby and Billy Play Ball Passage* (Appendix 5). The sentence "They take a bat, a ball and a glove" is not meaningful to Irish children, as a bat and glove are not used to play ball. If this sentence was adapted for Irish children, it would read "They take a sliothar and a hurling stick" (Sweeney et al., 1999).

Karnell (1995) stated that a possible weakness of nasalance measurements was that they did not distinguish between nasal acoustic energy that was due to nasal resonance and nasal acoustic energy due to turbulent nasal airflow. The *Zoo Passage* has been used routinely to obtain nasalance measurements and this passage has a high incidence of high pressure consonants. Karnell (1995) believed that, using this passage, it was not possible to determine the extent to which nasal acoustic energy was due to hypernasal resonance (occuring on vowels and semivowels) and audible turbulent nasal airflow (occuring on high pressure consonants). He hypothesised that a speech sample devoid of nasal consonants and oral pressure consonants (i.e. low pressure consonant sample) would enable measurement of nasal acoustic energy that is due to nasal resonance and not influenced by turbulent nasal airflow. Karnell (1985) compared nasalance scores on the high pressure (HP) speech stimulus (which contained stops, fricatives and affricatives) and on the low pressure (LP) speech stimulus (which contained vowels, glides and liquids). Using a nasalance cut-off value of 31%, Karnell divided all participants into three groups according to obtained scores. Group 1 had nasalance scores above 31% for the HP sentences and the LP sentences. This group was assumed to have velopharyngeal dysfunction. Group 2 had nasalance scores below 31% on HP and LP sentences. It was assumed that these speakers had normal velopharyngeal function, validated using perceptual assessments. Group 3 had mixed results (i.e. some had nasalance scores above 31% on HP sentences and below 31% on LP sentences, others had scores below 31% on HP sentences and above 31% on LP sentences). Karnell hypothesised that participants who had nasalance scores above 31% on HP sentences and below 31% on LP sentences may have had audible nasal turbulence. The presence of nasal turbulence may have caused an increase in nasalance scores. The main problem with this study was that no perceptual evaluations were made to validate the above groupings. If the participants who obtained nasalance scores above 31% on HP sentences and below 31% on LP sentences had perceptible nasal turbulence, Karnell's results would have been more conclusive. Nevertheless, this study was a useful contribution, as it highlighted the possible influence of nasal emission and nasal turbulence on nasometry results.

Indeed, in earlier studies the influence of nasal emission and turbulence on nasometric results were not considered. This fact may, according to Karnell (1995), explain the variability in agreement between nasalance measures and perceptual judgements found in earlier studies (Dalston et al., 1991b; Watterson et al., 1993). Karnell (1995; p. 9) states that " if the percept of audible nasal emission of air is lumped together with the percept of hypernasal resonance and, if the acoustic effects of turbulent nasal airflow has a proportionally greater or lesser impact on the nasalance measure compared to the perceptual rating, then the perceptual rating may disagree with the nasalance measure". This study highlights the importance of the speech samples used in nasometry research. Karnell suggested that low pressure and high pressure samples should be used if participants exhibit nasal airflow as well as resonance problems. Watterson et al. (1998a) reported that the Karnell (1995) study was of limited clinical value without perceptual validation. Watterson et al. (1998a) assessed twenty five children, five children had a history of no communication disorder, and twenty children were diagnosed as having nasal emission using a mirror fogging test. This involved placing a mirror under one nostril and occluding the other nostril while the child repeated single words and a short phrase. Perceptual ratings of nasality were made by seven experienced listeners. Results indicated no significant difference between mean nasalance scores for HP sentences (30.28%) and mean nasalance scores for the LP sentences (28.98%). Mean perceptual ratings of nasality for the two stimuli did not differ significantly. These results did not support Karnell's results. However, Watterson et al. (1998a) did not carry out any perceptual assessment of nasal emission/nasal turbulence. Some of the children included in this study may have had inaudible nasal emission which was detected by the mirror test, but would not have been detected perceptually. It would have been useful to assess children with audible nasal emission and/or nasal turbulence to ascertain if the nasalance scores were elevated for HP sentences compared to LP sentences. It is possible that perceptual judgements during production of HP sentences may also have been elevated when nasal emission was present.

It is evident from the literature that different speech stimuli provide different information regarding speech and velopharyngeal function. A speech stimulus devoid of nasal consonants will aid identification of patients with velopharyngeal dysfunction. Inclusion of HP and LP sentences in the speech stimulus may help identify patients with nasal emission and/or nasal turbulence. By including LP sentences, the literature suggests that the effects of nasal turbulence on reliability and validity of the Nasometer will be eliminated. In order to compare nasalance scores with conversational speech, a speech stimulus containing a representative sample of conversational speech would be useful. Finally, nasal sentences will help identify patients with upper airway problems affecting speech.

2. 6. 4 Normal Studies Using The Nasometer

Fletcher et al. (1989) assessed 117 children with normal speech from Alabama using the Nasometer. Participants repeated or read the three standard passages from the Nasometer package. Analysis indicated no significant age or gender effects on scores but significant effect for different speech stimuli. Similar results were reported by Watterson et al. (1996) using simplified passages. In the later study, 20 normal speaking children, aged 4;4 years to 6;4 years, recited three passages - the Turtle Passage (simplified version of the Zoo Passage), the Mouse Passage (simplified version of the Rainbow Passage) and the Zoo Passage. The passages were recorded on the Nasometer and on an audio tape recorder. Mean nasalance scores and standard deviations for each passage are presented in Table 2.1.

Table 2.1. Mean nasalance scores and standard deviations for normal American speaking children during the production of the five Nasometer passages (Fletcher et al. 1989; Watterson et al., 1996).

Stimulus	**Passage Type**	**Mean**	**SD**
Zoo Passage	devoid of nasals	15.53	4.86
Rainbow Passage	all consonant types	35.69	5.20
Nasal Passage	increased nasal consonants	61.06	6.94
Turtle Passage	devoid of nasal consonants	15.8	2.9
Mouse Passage	all consonant types	32.5	6.7

Seaver et al. (1991) reported similar nasalance scores on the Rainbow, Zoo and Nasal Passages for normal American adults. One hundred and forty eight participants, between the ages of 16 and 63 years, with normal speech patterns characteristic of four geographical regions were assessed using the Nasometer. They found differences in nasalance scores for different regions and suggested that regional norms needed to be established for clinical use of the Nasometer. This suggestion was later supported by a cross cultural study carried out by Dalston et al. (1993).

The Simplified Nasometric Assessment Procedures were used to assess normal American speakers in order to obtain norms for clinical assessments (MacKay & Kummer, 1994). Two hundred and forty six children were assessed during sentence production (4 sentences contained high pressure consonants, and 1 nasal sentence). Seventy six children were assessed reading 2 passages: Bobby & Billy which contained 8% nasals and A School Day which was devoid of nasals (Appendix 5). Mean nasalance scores and standard deviation scores for normal children were reported (Table 2.2).

Table 2.2. Mean and standard deviation scores for normal American children during the production of sentences and two passages.

Stimuli	Passage Type	Mean	SD
Sentences	devoid of nasal consonants	11.0 - 12.9	3.5 - 4.4
Nasal Sentence	increased nasal consonants	56.9	7.4
Bobby & Billy Passage	8% nasal consonants	15.4	2.8
A School Day	devoid of nasal consonants	10.3	3.1

Comparing MacKay and Kummer's nasalance scores for normal American children to the results of Fletcher et al. (1989) and Watterson (1996), it is apparent that MacKay and Kummer's scores are lower. On the sentences and the passage which are devoid of nasal consonants, the nasalance scores are approximately 5% lower. On the Bobby & Billy Passage (8% nasal consonants) the nasalance scores are the same as the nasalance scores on Zoo and Turtle Passage (devoid of nasal consonants) in the Watterson et al., study (1996). Regional variation in scores or the different speech stimuli may explain the difference in scores.

Nasalance scores for normal speakers in Australia were reported by van Doorn and Purcell (1998). Two hundred and forty five children who had normal speech were assessed using the Nasometer. Recordings were made during the production of the Zoo Passage (devoid of nasal consonants) and the Nasal Passage. Australian children were found to have slightly lower nasalance scores for the Zoo Passage (13%) and the Nasal Passage (59.6%) than American children (Fletcher et al., 1989; Watterson et al., 1996). Although the results from the studies compare well there is, according to van Doorn and Purcell (1998), a slight difference of approximately 2% for American and Australian children on the Zoo Passage. This, they stated, illustrates the dialectal difference between the two populations.

Nasalance scores in normal Finnish were reported by Haapanen (1991a). She assessed 50 participants between the ages of 3 and 52 years with normal speech. Each subject repeated 3 different sentences, one containing low pressure consonants and vowels, the second containing high pressure consonants and vowels and the third containing nasal consonants. Nasalance scores for each sentence type were reported (Table 2.3). Ninety percent of Finnish speakers in the study had a mean nasalance score below 20 -21% and the remaining 10% had scores between 22 and 29% indicating borderline nasal resonance. Haapanen (1991a) reported that cleft palate speakers with mild hypernasality scored between 22 and 29 on the Nasometer.

Table 2.3 Mean and standard deviation of nasalance scores during the production of high pressure consonant (HP), low presure consonant (LP) and Nasal sentences.

Stimulus	**Mean**	**SD**
HP sentence	14.2	4.6
LP sentence	11.1	5.1
Nasal sentence	69.2	7.5

Trindade et al. (1997) divided their participants according to age: children below 11 years; adolescents 11 to 17 years and adults. Each subject repeated 4 passages in Brazilian Portuguese, two Zoo type passages and two Nasal Passages. The first Zoo Passage (Zoo) was devoid of nasal consonants, while the second (Zoo2) was devoid of nasal and high pressure consonants. The first Nasal Passage contained 43% nasal consonants and the second (Nasal2) contained 66% nasal consonants. The results for each group and passage are presented in Table 2.4.

Table 2.4 Nasalance scores for normal Brazilian Portuguese speakers for 4 different phonemically balanced sentences

	Zoo		**Zoo2**		**Nasal**		**Nasal2**	
Age	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
< 11yrs	9	3.3	10	5.2	48	4.8	51	6.5
11- 17 yrs	12	6.9	15	8.5	50	6.2	52	5.8
> 17 yrs	13	5.2	15	6.2	47	6.2	47	6.8

Anderson (1996) assessed 40 adult Puerto Rican Spanish females with normal speech. Participants were aged between 21 and 43 years. They were recorded on the Nasometer during the production of five nasal sentences (each containing different proportions of nasal consonants) and two paragraphs (paragraph 1 contained nasal and non-nasal consonants representative of consonant distribution of Spanish, and paragraph 2 was devoid of nasal consonants). Mean nasalance scores, standard deviation and range of scores were reported (Table 2.5).

Table 2.5 Mean nasalance scores, standard deviation and range of scores for normal speaking females of Peurto Rican Spanish.

Stimulus	**Passage Type**	**Mean**	**SD**	**Range**
Paragraph 1	nasal + non-nasal consonants	36.2	7	23.3-49.6
Paragraph 2	deviod of nasal consonants	21.9	8.6	7.1-42.4
Nasal sentences	increased nasal consonants	62	7.7	37.5-78.7

Nasalance scores for paragraph 1 were comparible to the scores for American adults on the Rainbow Passage which also contains nasal and non-nasal consonants. However scores for paragraph 2 (no nasal consonants) were higher for the Peurto Rican Spanish-speaking females (21.9%) than for American adults during production of the Zoo Passage (16%). Anderson (1996) stated that the group mean for the paragraph 2 fell within the range of normal values of American female speakers as reported by Seaver et al. (1991). The language difference between studies may explain the difference in group means. In order to compare Anderson,s results with other studies, nasalance values for Puerto Rican males and children who speak Spanish are necessary.

Anderson (1996) included 5 nasal sentences in her study. The proportion of nasal consonants to the total number of consonants in the sentences varied from 33.3% to 60%. She found that, in general, higher nasalance scores were found for sentences with a higher proportion of nasal consonants, but that there was not a one-to-one correspondence between the proportion of nasal consonants and mean nasalance scores. The sentence containing the highest proportion of nasal consonants did not have the highest group mean. Trindade et al. (1997) reported a 1% increase in mean nasalance scores for a female Brazilian Portuguese-speaking population when the passage had an increase in proportion of nasal consonants (43% to 66%). The lack of one-to-one correspondence between the proportion of nasal consonants and mean nasalance scores may indicate that there is a critical proportion of nasal consonants. Once a passage contains a certain critical proportion of nasal consonants, the passage should be considered as a Nasal Passage for use in nasometry. Further investigation is required to ascertain a critical proportion of nasal consonants in the speech sample for nasometry.

2. 6. 4 (i) Influence of Age on Nasalance Scores in Normal Speakers

Haapanen (1991a) reported that on specific speech samples chronological age seemed to have an effect on nasalance scores. For high pressure consonant sentences, the mean and range of nasalance scores marginally decreased as age increased. However, on nasal sentences the range of nasalance scores increased with age.

Trindade et al. (1997) found that children had significantly lower scores on non-nasal passages than adults. They say that this result may be explained by the fact that children have smaller nasal cavities than adults. This, however, according to the authors, does not explain why children do not have lower nasalance scores on the nasal passages. Results of American studies do not support the assertion that there is a difference between nasalance scores for adults and children (Fletcher et al., 1989; Seaver et al., 1991).

Age difference did not significantly influence nasalance scores in a study on normal Australian children (van Doorn & Purcell, 1998). However, in this study the participants were aged between 4 and 9 years and age differences were minimal. It would be useful to investigate normal nasalance scores for Australian adults to allow for comparison of scores for different age groups.

2. 6. 4 (ii) Influence of Gender on Nasalance Scores in Normal Speakers

Gender difference in mean nasalance scores was reported by Seaver et al. (1991). Female participants were found to have significantly higher scores than male participants on the Nasal Sentences. Trindade et al. (1997) suggest that, despite being statistically significant, it is questionable whether the gender difference is clinically significant, as nasalance score differences across gender averaged only 2 scale points. No statistical difference was found between male and female nasalance scores in studies by Trindade et al. (1997) and van Doorn and Purcell (1998). Both these studies included children and adults.

2. 6. 4 (iii) Conclusion from Normal Studies

Normal studies using the Nasometer have indicated variation in normal scores due to influences of language, dialect, populations studied and the speech stimulus used. Comparison of all studies is difficult due to the fact that different languages were used in these studies. However, the literature illustrates that language variation results in variation of normal nasalance scores and therefore normative scores are required for each language. Comparing English-speakers of different dialects has been found to produce varying normal nasalance scores (Seaver et al., 1991).

The use of different populations makes comparisons and conclusions difficult. Normal scores on adults and children are required for each population, and gender differences for the populations should be evaluated. Research results to date are inconclusive.

Unfortunately, many studies did not report normal nasalance scores for passages containing the normal distribution of consonants (Haapanen, 1991; Trindade et al., 1997; van Doorn & Purcell, 1998). Thus comparison with other studies which include the Rainbow Passage (phonemically balanced) was difficult. The Zoo Passage (devoid of nasal consonants) has been found to be most suitable for accurate detection of hypernasality (Dalston et al., 1991b; Watterson et al., 1993), while the Nasal Passage has been found to be accurate for detecting hyponasality (Dalston et al., 1991c). As a result, these two passages or similar passages have been used in many normal studies. However, if the Nasometer is to be used for clinical evaluation of speech, a speech sample that represents conversational speech also needs to be used.

The need to establish nasalance scores for normal speakers using the language, dialect and age of the clinical population in which the Nasometer is to be used is underscored by the above studies.

2. 6. 5 Comparison between Nasalance Results and Other Assessment Results for Normal Speakers.

Dalston (1989) reported a study using simultaneous photodetection and nasometry to monitor velopharyngeal behaviour. He assessed six normal adults as they produced six repetitions of 10 sentences. A photodetector probe was placed approximately 5mm below the resting level of the velum. The Nasometer headset was then placed in position. Results indicated that Nasometer and photodetector signal maxima were within 30 milliseconds of one another, indicating a correlation between velopharyngeal opening and nasal acoustic energy. The study suggests that during the production of nasal consonants by normal speakers, there is a close temporal correspondence between the output signals of the two different techniques. Further investigation is required to evaluate the relationship between these techniques for speakers with velopharyngeal inadequacy.

Nasalance scores for normal African-Americans and white-Americans were obtained and compared to nasal cross-sectional area measurements for the same group (Mayo, Floyd, Warren, Dalston & Mayo, 1996). Nasal cross-sectional area was measured using the aerodynamic assessment technique (PERCI) originally described by Warren and Dubois (1964). Results indicated no significant difference between groups for nasalance scores on the Zoo Passage; however, white speakers had significantly higher nasalance scores on the Nasal Passage. No racial difference in nasal cross-sectional area was found. The correlation between measurements was poor. This study ruled out anatomical difference in races as a possible cause for difference in nasalance scores for the Nasal Passage.

2. 6. 6 Reliability of the Nasometer

Test-retest reliability of the Nasometer was examined in the study by Seaver et al. (1991), where normal nasalance scores were obtained for adults from four geographical regions in America. Forty participants read the Zoo, Rainbow and Nasal Passage three times in succession. Analysis of the data indicated that 97% of the mean nasalance scores for any reading of the Zoo Passage were within 3 percentage points of the score on any other reading of that passage. The largest difference in scores between readings was 6%. Ninety one percent of the mean nasalance scores were within 3 percentage of points for other readings of the Rainbow Passage, while ninety four percent of the mean nasalance scores were within 3 percentage points for other readings of the Nasal Passages. On the Nasal Passage the largest variation between scores was 8%. Results from a study by Mayo et al. (1996) support these findings.

Van Doorn and Purcell (1992) reported that on a second reading of the passages, 100% of readings for the Zoo Passage had mean nasalance scores that were within four nasalance points and 92% of mean nasalance scores were within five nasalance points on the Nasal Passage. On the Nasal Passage, the greatest difference between reading was 9%. However, they also reported considerable variability in repeated nasalance measures during repetition of the passages across different assessment sessions. Variability within the session was less substantial than variability across sessions.

The above results indicate that there is variation in nasalance scores for individuals during repetition of the same speech sample but that in most cases the variation is not greater than 4% (Trindade et al. 1997). Seaver et al. (1991) reported greater variation on repetition of the Rainbow Passage (phonemically balanced) than on repetition of the Zoo

Passage (devoid of nasal consonants). This may be explained by the phonetic content of the Rainbow Passage. Van Doorn and Purcell (1998) also report greater variation in children than in adults. In both cases the increase in variation may be due to the timing of velopharyngeal opening and closing that is required. In the Zoo Passage, velopharyngeal closure is sustained, whereas in the Rainbow Passage the timing of velopharyngeal movement may vary. Less precise timing of velopharyngeal closure in children compared to adults may be responsible for increased variation in nasalance scores when the passage contains nasal consonants (van Doorn & Purcell, 1998).

Nichols (1999) reported that, using sets of 10 sentences for non-nasal and nasal speech stimuli, a high level of reliability was found for mean nasalance scores (Cronbach alpha coefficient of 0.95 for non-nasal sentences and 0.94 for nasal sentences). He also reported that the reliability of the mean nasalance scores decreased with a reduction in the number of sentences in each speech category. He stated that clinical reliability of a short sentence list is questionable. However, Nichols (1999) pointed out that, in a study with many participants, the overall means may be reliable even when individual assessments are not.

Contrary to Nichols' (1999) findings, Watterson, Lewis and Foley-Homan (1999) found high levels of correlations between longer speech stimuli (44 syllable sentences) and short speech stimuli (17 syllable and 6 syllable sentences).The Pearson correlations were 0.95 for the 44 syllable sentences and the 17 syllable sentences, and 0.93 for the 44 syllable sentences and the 6 syllable sentences. This indicated a high criterion validity for short speech stimuli. Unfortunately, Watterson et al. (1999) did not examine the test-retest reliability of the short speech stimuli.

2. 6. 7 Clinical Studies Using Nasometry

2. 6. 7 (i) Validity of the Nasometer

Validity of the Nasometer has been established by evaluating correlations between nasalance scores and perceptual assessments. Studies have been carried out to examine the extent to which measures of nasalance correspond to perceptual judgments of nasality. In the studies, perceptual judgments of nasality were made by one listener (Dalston et al., 1991b) or by a panel of listeners (Paynter et al., 1991) using various rating scales. Rating scales used included: a 9 point scale, where 1 indicated normal resonance, +5 indicated severe hypernasality and -5 indicated severe hyponasality (Paynter et al., 1991); a 6 point equidistant scale (Dalston et al.,1991b); and two six point equidistant scales, one for hypernasality and one for hyponasality (Nellis et al., 1992). Perceptual judgments of nasality were made during the production of the standard nasometry passages (Paynter et al., 1991) and modified passages (Watterson et al., 1996). In some studies, simultaneous audio recordings were made during nasometry assessments. In the study by Dalston et al. (1991b) perceptual judgments of nasality were made during clinical assessments using different speech samples, while nasometric analyses were made during production of the Zoo Passage. Correlation coefficients were calculated to assess the relationship between perceptual judgments of nasality and nasalance scores (Table 2.6). There was considerable variation in results.

Table 2.6 Studies reporting correlation coefficients for perceptual judgements of nasality and nasalance scores during the production of different speech stimuli.

Study	Speech Stimulus	Correlation Coefficient
Paynter et al., 1991	Zoo & Nasal Passages Rainbow Passage	0.66 0.63
Dalston et al., 1991b	Zoo Passage	0.82
Haapanen 1991b	Nonsense syllables + 1 word containing a nasal + HP + LP sentences	0.75 - 0.79
Dalston & Seaver, 1992	Rainbow Passage	0.63
Nellis et al., 1992	7 sentences	.02 - 0.43 (hypernasality) .05 - .61 (hyponasality)
Watterson et al., 1993	Modified Passages (devoid of nasal consonants)	0.49
Watterson et al., 1996	Zoo Passage Turtle Passage (simplified Zoo Passage) Mouse (simplified Rainbow Passage	0.70 0.51 0.32

Haapanen (1991b) reported good correlations between nasalance scores and her perceptual 'index of hypernasality'. The index involved opening and closing the nares while the child repeated nonsense syllabes and sentences. The hypernasality index was determined by counting the number of words perceived to have shifted in tone.

In the Dalston et al. (1991b) study, only one listener was used to judge nasality in the study and intra-judge and inter-judge reliability were not assessed. One other problem with this study was that patients were assessed using different speech samples in each technique, ranging from a single word utterance to a complex passage. In this study no reference was made to the degree of hypernasality and the differences in nasalance scores.

The Nellis et al. (1992) study reported lower correlations between ratings of hypernasality and mean nasalance scores compared to the previous studies. Results indicated that the mean judges' rating of hypernasality did not increase systematically with increasing nasalance scores. On nasal loaded sentences, mean judges' ratings of hyponasality decreased as nasalance scores increased. Results also indicated that nasalance scores, ratings of hypernasality and ratings of hyponasality differed by sentence as a result of phonetic context. The authors state that one possible reason for the poor correlations found in this study was that judges were asked to rate each sentence for both hypernasality and hyponasality. Understandably, this may have caused difficulty for the judges, who felt that speakers had to be rated on both scales even if, for example, hyponasality was not present. Results may have been more reliable if judges were told to rate on either one scale or the other.

In the two Watterson et al. studies (1993; 1996) different correlation coefficients were reported. Unfortunately, Watterson et al. (1996) did not comment on the different correlations in the two studies. It may be due to the fact that different speech stimuli were used in each study. A low correlation ($r = 0.32$) between nasalance scores and perceptual judgements was found in Watterson et al.'s (1996) study on the Mouse Passage which contained 11% nasal consonants. This low correlation was explained by the authors by the fact that a passage containing nasal consonants had previously been found to be unhelpful in identifying hypernasal patients by nasalance measures (Watterson et al., 1993). It should be noted, however, that assessment of nasal resonance in speech does not and should not limit itself to a clinical assessment. Therapists need to ascertain what everyday speech is like in order to make clinically appropriate judgements regarding management. A speech sample used for assessment should reflect normal conversational speech, which includes nasal consonants.

Paynter et al. (1991) examined the degree to which the mean nasalance scores agreed with the listener ratings in classifying each speech sample as being normal, hypernasal or hyponasal. On the Nasal Passage, 48% of all categorizations based on nasometry agreed with perceptual judgements. On the Zoo Passage the agreement was 66% and on the Rainbow Passage agreement was 67%. The findings suggest that when using three broad categories of hypo-, normal, and hyper-, the nasalance scores from the three different passages did not categorize the participants in a manner similar to the listeners. The authors suggest that the passages are better at categorising if each passage is used to identify a separate category (i.e. the Zoo Passage would be the stimulus choice for identifying hypernasality and the Nasal Passage for identifying hyponasality). This study highlights the importance of the speech sample used in studies evaluating the usefulness of the Nasometer.

Correlations between perceptual ratings of nasality and nasalance scores reported in the literature varied from 0.02 (Nellis et al., 1992) to 0.82 (Dalston et al., 1991b). The variation in correlations have been reported to be due partly to the speech stimuli used in the various studies. Overall, relationships improve if ratings of hypernasality are correlated with passages devoid of nasal consonants and if ratings of hyponasality are related to nasal speech passages. Variations may also be due to the presenting problems of the participants. If participants with nasal emission/turbulence were included in the study, the nasalance scores may have been inflated (Karnell, 1995). Hence, the relationship between perceptual and nasometric measurements may have been influenced. Watterson et al. (1998b) reported sources of error in nasality ratings which underestimate the association with nasalance scores. They found that when a mean perceptual rating from a panel of raters was compared with nasalance scores the correlation between the assessments was weak. Correlations improved when a single expert rater was used to rate nasality and this rating was compared with nasalance scores. Watterson et al. (1998b) also found that using a restricted range of nasality ratings for perceptual assessments resulted in weak correlations.

2. 6. 7 (ii) Cut-off scores

Varying cut-off values have been reported to distinguish between normal nasalance and abnormal nasalance in the literature. Dalston et al., (1991b) found a cut-off value of 32% for the Zoo Passage to be best for the detection of hypernasality; however, Watterson et al. (1993) used a cut-off value of 26%. These differences are difficult to interpret because of the different methodologies used in each study (van Doorn & Purcell, 1998).

Dalston et al. (1993) carried out a cross-dialect and cross-cultural study on nasometric sensitivity and specificity. Prediction analyses revealed that maximum efficiency (i.e. highest sensitivity and specificity scores) was obtained using different cut-off nasalance values for each of three patient groups (two clinics in the US and one in Spain). This indicated that nasalance scores need to be obtained for different regions so that dialect norms can be calculated and compared to a clinical population in order to obtain cut-off values. Unfortunately, the speech samples and the perceptual scales used varied from one clinic to another. This made comparison of the relationships between perceptual ratings and nasalance scores for each clinic difficult.

2. 6. 7 (iii) Sensitivity and Specificity

Dalston et al. (1991b) reported on the extent to which the Nasometer correctly identified patients who had previously been categorised as hypernasal on perceptual judgements. They reported that, using a cut-off value of 32% nasalance, the Nasometer had a test sensitivity of 0.89 and a specificity of 0.95. Using the same cut-off value, Hardin et al. (1992) reported weaker associations between listener ratings of hypernasality and nasalance scores. They found a sensitivity score of 0.57 and a specificity score of 0.91. Using a lower cut-off value of 26%, Hardin et al. (1992) found an improved sensitivity of 0.76 and specificity of 0.85. Hardin et al. (1992) also reported that, by eliminating participants who had undergone pharyngeal flap surgery from the study, sensitivity (using a cut-off value of 26%) increased to 0.87 and specificity improved to 0.93. Variations in sensitivity and specificity may be due to the cut-off values used, and the populations assessed. Hardin et al. (1992) stated that the large discrepancies in associations between the two assessments are difficult to explain, but possible factors included the difference in the number of judges in each study, the different levels of experience of the judges in rating hypernasality and the different speech samples used for perceptual ratings.

Nasometer test sensitivity and test specificity were determined for the patient group in the study by Watterson et al. (1993). Using a cut-off value of 22% a sensitivity for the Zoo and Turtle Passages was 0.77. The specificity was lower for the Zoo Passage (0.50) and the Turtle Passage (0). This low specificity was explained by the authors by the fact that only 2 participants were judged to have normal nasality.

Sensitivity and specificity for HP sentences and LP sentences was calculated by Watterson et al. (1998a). They used a cut-off value of 26% for HP and LP sentences. No significant differences between the two stimuli were found and therefore sensitivity and specificity for pooled data of HP and LP sentences were reported. Sensitivity was 0.84 and specificity was 0.88. The results of this study indicated an improved sensitivity and specificity compared to earlier studies by Watterson et al. (1993). The authors state that this may be due to methodological differences in the studies. In the latter study listeners were informed that a rating of '0', '1' or '2' indicated normal nasality, whereas in the earlier study a post-hoc decision was made that ratings of '1' or '2' would be considered normal. Thus, in the latter study the listener knew that a rating of '2' was considered normal, but in earlier studies, as far as the listener was aware, it may have represented mild hypernasality.

Dalston et al. (1991a) carried out a preliminary investigation concerning the use of nasometry in identifying patients with hyponasality and/or nasal airway impairment. They evaluated the sensitivity and specificity of the Nasometer in correctly identifying the presence or absence of hyponasality. Test sensitivity was reported to be 0.48 and test specificity was reported to be 0.79. However they found that when patients with audible nasal emission were excluded from the analysis, the sensitivity rose to 1.0 and specificity rose to 0.85. No intra-judge or inter-judge reliability studies were carried out for perceptual judgements. Hardin et al. (1992) reported a similar sensitivity score of 1 and specificity of 0.87. They also excluded participants with audible nasal emission.

The literature reports varying cut-off values for different studies. This may be explained by the use of normal nasalance data to indicate cut-off values. Thus, variations in normal scores will produce variations in cut-off values. To date, sensitivity and specificity results are encouraging and hence, this may be a more appropriate way of determining the value of the Nasometer as a clinical assessment tool than reporting correlation analyses (Dalston et al., 1993).

2. 6. 7 (iv) Standard Deviations in Nasometric Scores

Vallino-Napoli and Montgomery (1997) examined the use of standard deviation in nasometric results. Results indicated that standard deviations on their own were unable to distinguish between different severity groups of hypernasality (i.e. mild hypernasality, moderate/severe hypernasality). Vallino-Napoli and Montgomery concluded that standard deviation scores on their own had limited clinical value.

2. 6. 7 (v) Nasometry and Nasal Obstruction

Parker, Maw and Szallasi (1989) studied the use of the Nasometer in the selection of patients for adenoidectomy. Three normal adults were assessed on the Nasometer with nostrils open and again with the nostrils occluded. Each subject repeated phonemically balanced test phrases and the Zoo Passage. They found a significant fall in nasalance scores on test phrases when the nostrils were occluded. Further research in this area is recommended by the authors, who claim that nasometry may aid in selection of patients for adenoidectomy.

2. 6. 8 Nasometry Conclusions

All studies agree that the Nasometer can be a useful tool in evaluating nasality in speech when used with other forms of assessment. It is also agreed that nasalance scores are required for normal speakers representative of the clinical population if scores are to be meaningful (Seaver et al., 1991; Trindade et al., 1997).

There is still much controversy regarding the level of correlation between nasalance scores and perceptual judgements of nasality. In many of the studies there were substantially different speech samples used for the perceptual judgements and the Nasometer assessments. Also, ratings of nasality in all these studies have been undertaken using equidistant scales with all their disadvantages. The use of more descriptive scales of nasality may improve correlations. Use of test sensitivity and test specificity should be included in the evaluation of the relationship between nasometry and other assessment techniques. This approach provides more detailed information regarding the agreements and disagreements in classification of normal or abnormal speakers.

There is disagreement in the literature regarding the value of controls and normative data. In many studies different cut-off nasalance values were used to distinguish between normal and hypernasal speakers. Cut-off values in the various studies differed as much as 8% nasalance score. Watterson et al. (1996) used a cut-off score between normal and hypernasal children of 22%, whereas Karnell (1995) used a cut-off score of 30%. This makes comparison of studies difficult. The use of mean nasalance scores for normal speakers combined with clinical data may help identify reliable cut-off values for research and clinical use.

Although controversial, nasal emission and nasal turbulence have been shown to influence nasometry results and therefore in future studies nasal airflow must be taken into account when evaluating research results (Karnell, 1995). The use of low pressure consonant sentences and high pressure consonant sentences as speech stimuli for the Nasometer may help distinguish between hypernasality and nasal airflow errors.

This review of the nasometry studies indicates that the standard passages (Zoo, Rainbow and Nasal) have been used extensively in research. These passages have problems - some are too complex for children (Rainbow Passage, Watterson et al., 1993); others do not represent normal conversational speech (Zoo Passage and Nasal Passage). It has been agreed that the Zoo Passage is useful in identifying hypernasality (Dalston et al., 1991b; Vallino- Napoli & Montgomery, 1997) and that the Nasal Passage is useful in identifying hyponasality (Dalston & Seaver, 1992; Watterson et al., 1993; Vallino-Napoli & Montgomery, 1997). However, if comparisons between perceptual judgements of nasality and nasalance scores are to be made, a speech sample which represents conversational speech should be included. If further comparison of high pressure consonant sentences and low pressure consonant sentences are to be made, these sentence categories should be distinguished in the sample. Hence, a speech sample that represents normal conversational speech should be used for nasometric assessment. This could be presented in a manner that allows each sentence category to be distinguished from each other and assessed independently, i.e., high pressure consonants, low pressure consonants (devoid of nasal consonants), mixed consonants (all consonant types) and nasal sentences (containing high proportion of nasal consonants). It is also recommended that the Nasometry speech sample be included in the speech sample for perceptual assessment, thus reducing the number of variables which influence the relationship between the two measurements.

2. 7 Aerodynamic Assessment

Nasal emission of air during speech production is a problem which can be associated with repaired cleft palate and non cleft velopharyngeal dysfunction. Various simple devices have been developed to indicate the presence of nasal emission of air. The mirror test involves placing a mirror under the nostril and asking the child to repeat a syllable or word. The mirror fogs when it contacts with humidified air, indicating emission of air through the nose. The 'See-Scape' is an equally crude airflow detection device (Moon, 1993). This simple device consists of a flexible air tube which is positioned in the patients' nares. The tube is connected to a rigid tube, housing a small piece of styrofoam. If air leaks through the nares, the airflow displaces the styrofoam float upwards. These simple devices indicate the presence or absence of nasal airflow errors. They do not provide information regarding the amount of airflow (Moon, 1993). Warren (1975) stated that these basic tools differentiate between gross palate defects and normal palatal function, but they are unreliable in differentiating between slight or moderate palatal incompetence and normal function. He stressed that, where the speech symptoms vary from one speech situation to another or where compensatory articulation exists, the devices are unreliable. More elaborate instrumental assessment tools are required to measure the amount of nasal airflow and to identify where the leakage of air is occurring (Moon, 1993). These include aerodynamic measurement systems which record airflow volumes and pressure flows during speech. Two types of flow meters have been used in devices during the last three decades; one is the anemometer and the other is the pneumotachograph. In this section, aerodynamic systems will be described and different techniques will be evaluated. Aerodynamic studies of normal speakers, the speech stimuli used in studies, and studies of their relationship with other assessment procedures will be reviewed. The need for further research in aerodynamic assessments will be highlighted.

2. 7. 1 Anemometry

Airflow may be recorded by measuring the changing pattern in temperature of the air as it is inhaled and exhaled. Various electrical methods of temperature measurements have been devised using the thermal effect on electrical resistance (warm wire anemometry and a thermistor), the influence of temperature on capacitance in semiconductors and thermally induced voltages (O'Neill & Malone, 1977). The warm-wire anemometer and the thermistor are more sensitive and hence, much of the work on airflow measurement has been limited to these temperature transducers.

The warm-wire anemometer is a thin wire heated by an electrical current passing through it (Figure 2.2). Air passing over the wire cools it and the electrical current is altered. The thermistor is a hard ceramic-like device composed of a compressed mixture of metallic oxides, moulded into a bead, rod or disk (Geddes & Baker, 1989). The resistance of the metal oxides decreases with increases in temperature and the thermistor is usually connected directly to a monitoring instrument (O'Neill & Malone, 1977). However, the warm-wire anemometer is sensitive to variations in room temperature and humidity, resulting in unstable baselines on the graph (Painter, 1979).

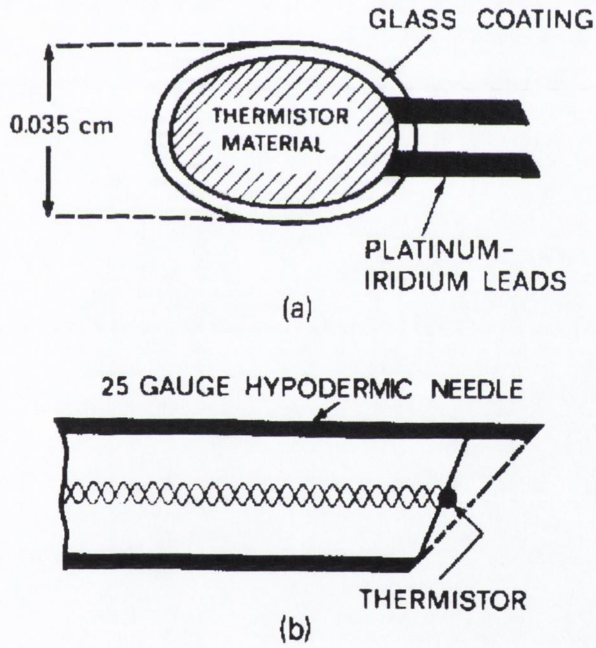


Figure 2.2 Diagrammatic representation of two thermistors used in a warm-wire anemometer (O'Neill & Malone, 1977).

The Exeter Nasal Anemometry System which utilised a thermistor was devised by Ellis et al. (1978). This simple device allowed for assessment of nasal airflow errors during production of isolated words. A small hand-held under-nose mask contained a thermistor which detected changes in air temperature. The varying electrical resistance yielded by the thermistor was input into the anemometer control unit. A microphone was placed in front of the mouth. The two signals could be recorded on an audio tape and sent to a laboratory for analysis, or they could be computed on a BBC computer and displayed graphically on the screen (Figure 2.3).

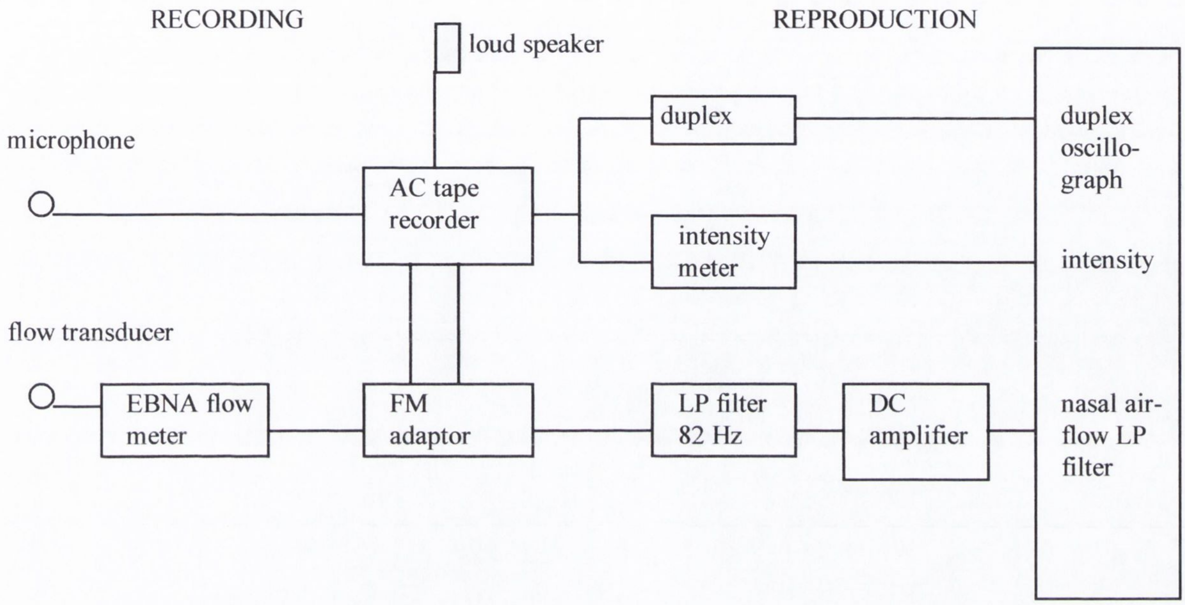


Figure 2.3 Exeter Nasal Anemometry System, indicating laboratory and computer-based system. Adapted from Hutters and Brondsted (1992).

Hutters and Brondsted (1992) evaluated the clinical usefulness of the Exeter Nasal Anemometer in the assessment of speech for clinical purposes. They outlined the major shortcomings of the Exeter Nasal Anemometry system as follows:

- the flow meter provided non linear recordings and therefore calibration was unreliable;
- the frequency response of the thermistor was poor, hence airflow during the production of words and sentences could not be assessed reliably;
- it did not distinguish between ingressive and egressive airflow;
- the temperature dependency of the meter may have artefactual influences and obtaining a zero baseline could be difficult.

Hutters and Brondsted (1992) found that reliable nasal airflow data could be obtained only for sounds in isolation. The device had limitations also as a tool in providing information about the dynamic airflow patterns during speech. They also reported difficulty in calibration. Warren (1975) stated that the meter, which measures nasal emission only, does not give an indication of how well the velopharyngeal sphincter functions. Factors such as respiratory effort and articulatory proficiency need to be taken into account when assessing speech and velopharyngeal function.

Despite the many shortcomings of the Exeter Nasal Anemometer, it has been used extensively in the clinical setting in the United Kingdom and Ireland, because it was accessible, inexpensive, easy to use and thought to provide a measure of nasal airflow (Hutters & Brondsted, 1992; Ellis, 1995, personal communication).

The Nasal Oral Ratio System (NORS) was developed at the University of Kent, Canterbury by Mirlohi et al. (1994). The system was based on the earlier Exeter Anemometry System and was developed to measure nasal airflow during speech. It consisted of a mask which was divided into a nasal and an oral section, each containing a thermistor (heat sensor) and a microphone (Figure 2.4). The mask was placed on the participant's face. During speech production the airflow and sound signals from the thermistors and microphones were recorded, amplified and passed on to a PC data acquisition board (Figure 2.5). The software program collected the data and performed analyses of oral airflow and nasal airflow. A ratio {of the differential (nasal-oral) airflow to the total (nasal+oral)} airflow was calculated. This ratio proposed to eliminate the effect of respiration on airflow measures. The ratio provided a value between -1 and +1, where +1 equalled complete nasal flow and -1 equalled complete oral flow.

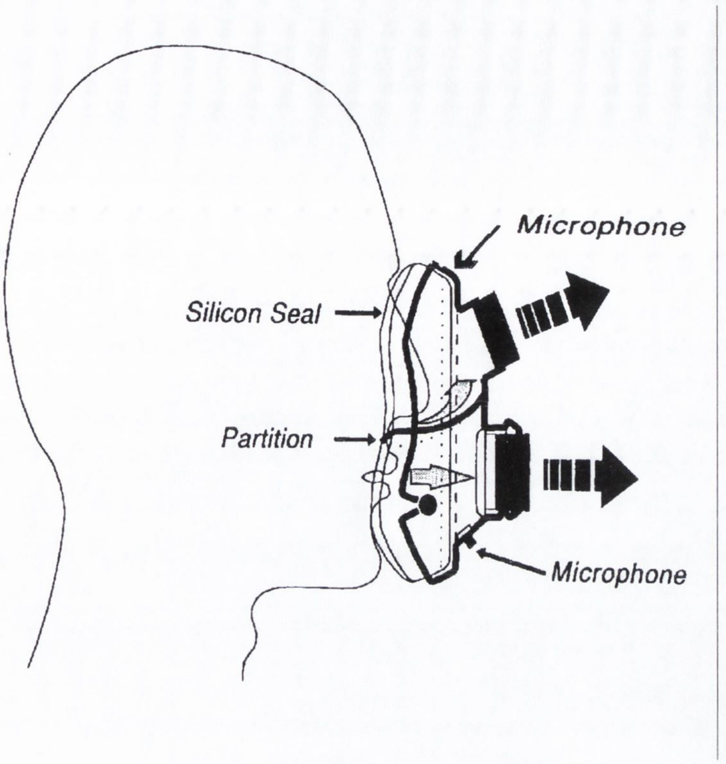


Figure 2.4 NORS mask, adapted from Mirlohi et al. (1994)

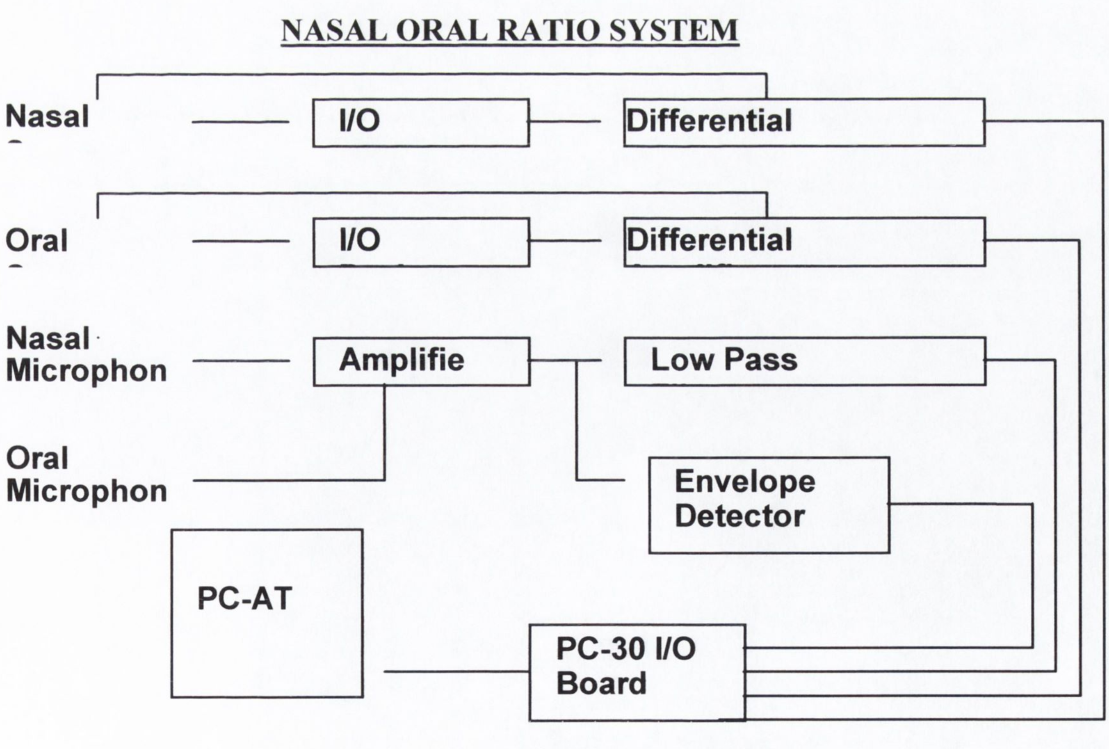


Figure 2.5 NORS Block Diagram, adapted from Mirlohi et al. (1994)

In a pilot study assessing normal speakers on the NORS, 20 samples of normal speech were analysed in an attempt to obtain normal airflow scores and graphs. Children between the ages of 7 and 10 years were recorded using 4 different speech samples - words, sentence repetition, automatic speech and conversational type speech. Significant problems were found with the system. The baseline on the graph was unstable, providing an unreliable baseline. The baseline measurement was affected by any movement of the mask and exhalation. The response time of the thermistor was slow and, as a result, there was a time delay between the speech output and the airflow output. This made it difficult to measure airflow for particular sounds. The quantitative scores provided were indefinite (i.e. normal oral sound values for the same sound ranged from +0.23 to -0.62 when the score should have been close to –1). The quality of the acoustic graph was poor and identification of any particular sound from the graph was difficult; hence, measurement of peak airflow values was unreliable (Sweeney, Sell and Grunwell, 1996).

The NORS was further developed and the Super Nasal Oral Ratiometry System (SNORS) (Main, Kelly and Manley, 1999) is now available. The system uses the same principles as the NORS, but it uses high-speed sensors to measure changes in airflow. Oral and nasal airflow are measured and a ratio score calculated. Main et al. (1999) refer to this ratio as a nasalance value. This, however, results in considerable confusion (Sell & Sweeney, in press). In previous studies nasalance refers to acoustic energy (Fletcher et al., 1989; Dalston & Warren, 1986; Haapanen, 1991a; Dalston & Seaver, 1992; Nichols, 1999); however, in the Main et al. (1999) study nasalance refers to airflow. The quantitative value using an airflow ratio needs to be validated. No normative data is available using the SNORS. Although the sensors are reported to have rapid reaction times, sharp increases in airflow graphs are not easily distinguished.

Because of the problems of heat sensors used in the above systems, it is now recommended that sensors with rapid reaction time be used to assess airflow in speech (Ellis, 1995, personal communication).

2. 7. 2 Pneumotachograph

The pneumotachograph measures differential pressures across a membrane in a tube. Volume of air and other facets of airflow can be determined from the pressure differential and dimensions of the tube. The pressure difference across the membrane is transmitted to a pressure transducer. This compares the two pressures with a known resistance, calculates the airflow and converts it into voltages which are passed on to an amplifier, analyser and recorder. This device has better frequency response time than the warm-wire anemometry system and the thermistor. It is also sensitive to direction of airflow. It allows for assessment and comparison of oral and nasal airflow during connected speech (Counihan, 1971). For the diagnosis of and research into cleft palate speech, Warren (1975) recommends a compromise between simple manometric devices and sophisticated research tools. He points out that the best compromise in terms of expense and simplicity is to obtain an instrument which could record a ratio of oral/nasal pressures during plosive consonant production in conjunction with nasal airflow measurements.

2. 7. 2 (i) Pressure/Flow Systems

Warren and DuBois (1964) described a pressure/flow technique for measuring the velopharyngeal orifice area during continuous speech. They based their method on a modification of the theoretical hydraulic principle. This principle assumes that the area of an orifice can be determined if the differential pressure across the orifice is measured simultaneously with the rate of flow through it. Thus, the velopharyngeal orifice area equals the following equation:

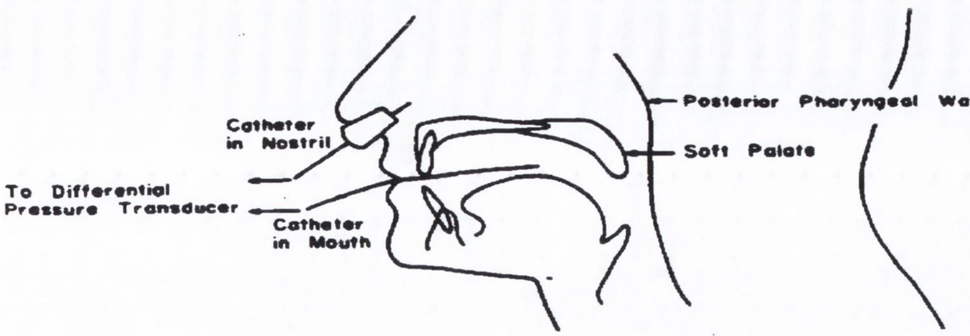
$$\text{Orifice Area} = \frac{\text{Rate of Airflow through Orifice}}{\rho \; 2\left(\frac{\text{Orifice Differential Pressure}}{\text{Density of Air}}\right)}$$

The advantage of this method is that parameters related to velopharyngeal closure (orifice size, oral pharyngeal pressure, orifice airflow and acoustic characteristics) can be assessed and evaluated simultaneously.

Warren (1967) used this pressure/flow system to assess nasal escape of air and velopharyngeal function. One catheter was placed in the left nostril and another in the oral cavity in order to measure the pressure drop across the orifice. The nasal catheter was secured by a cork which blocked the nostril, creating a stagnant column of air. Both catheters measured static air pressure and transmitted these pressures to a differential pressure transducer. A heated pneumotachograph connected to a plastic tube was placed in the participants' right nostril to measure nasal airflow. The parameters of pressure and airflow were converted to electrical voltages, amplified and recorded (Figure 2.6). The pressure/flow system, which was designed by Warren (1967), provided different measurements of pressure and airflow. The basic measurements included nasal airflow in millilitres per second (mls/sec) and oral and nasal pressure in centimetres of water (cm H_2O). These measurements were then used to calculate differential pressure (i.e. difference between pressure in the nose and pressure in the mouth) and velopharyngeal port area (cm^2).

2. 7. 2 (ii) PERCI (Palatal Efficiency Rating Computed Instantaneously)

Warren (1979) described the PERCI method for rating palatal efficiency. PERCI was a clinical tool, which was a development of Warren's (1967) system (Figure 2.6). He stated that the validity of the instrument depended upon its ability to separate the effects of velopharyngeal function from the activities of other speech structures. This approach was designed to record the difference between oral and nasal pressure. As in the earlier studies, airflow was used to calculate velopharyngeal orifice size.



Placement of pressure catheters

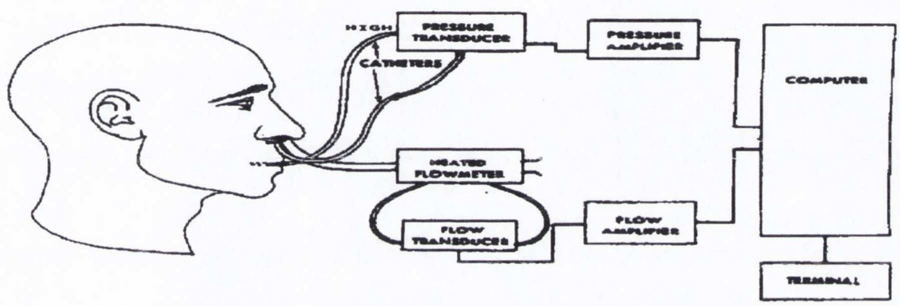


Figure. 2.6 Diagrammatic representation of tube placement for pressure measurements. Block diagram of pressure/flow system. Taken from the PERCI SARS Manual (1994).

Warren (1979) explained how the pressure/flow system works - closure of the velopharyngeal sphincter creates a pressure difference between the nose and the mouth. When complete velopharyngeal closure occurs, as during the consonant /p/, pressure in the mouth is determined by respiratory effort and will vary from about 3 cm H_2O to 7 cm H_20. Pressure in the nose will be zero since no air leaks into the nasal chamber. If there is velopharyngeal opening, the difference in pressure will vary with the size of the opening. Since the difference in oral and nasal pressures is used, the effects of respiratory effort can be cancelled out when the orifice is not completely closed.

2. 7. 2 (iii) PERCI SARS

The SARS (Speech Aero-Dynamic Research System) is the most recent development of the PERCI System described by Warren (1979). SARS is a computer software/hardware interface for in-depth assessment of speech aeromechanics and nasal airway patency (PERCI SARS Version 2.1 Manual, 1997). This computer-based system allows for measurements of:

1. oral pressure
2. nasal pressure
3. oral flow
4. nasal flow

Using these measurements, differential pressure and velopharyngeal port area can be calculated using the SARS software. In addition the system assesses:

nasal areas
nasal resistance
laryngeal resistance
voice analysis - sound pressure level, pitch, shimmer, jitter

Riski, Warren, Zajac and Lutz (1995) described the PERCI SARS aerodynamic assessment of speech. They pointed out the SARS system offers users greater control of their pressure/flow analysis with many options for measurements of pressure/flow and voice. The high speech data acquisition channel for voice has been added to the SARS. This has been found to be a significant advantage to the system as it allows simultaneous voice playback during analysis of the graphs. This improves validity of the system by providing the assessor with the acoustic signal of the speech sample as well as the pressure/flow graphs and measurements (Sweeney et al., 1996).

2. 7. 2 (iv) Gaeltec System

Another pneumotach system was described by Anthony (1980). His system was designed to allow for easy comparison of one physical quantity with another and to allow for adequate segmentation of the utterance. Various traces can be produced graphically showing paper speed, nasal flow, oral flow, laryngeal trace (indicating voicing) and a fundamental frequency trace (indicating frequency of vocal cord vibration or pitch). Anaesthetic masks are modified to hold separate pneumotach flow heads in a nasal section and an oral section. A rubber divider piece fits above the upper lip to separate the oral and nasal flows. During speech production the pressure drop accross the wire gauze in the pneumotach flow head is measured in millilitres per second. Data from the various traces allow for segmentation of the utterances into phonemic units and syllables. However, Anthony (1980) points out that this segmentation is artificial and to some extent arbitrary, but necessary in order to measure the aerodynamics of speech. Anthony (1980) states that the data provided by the Gaeltec system provides an insight into the dynamic organisation of speech production and indicates problems that can arise if there is some anatomical abnormality. One drawback of the system however, is that the audio recording cannot be made simultaneously with the airflow assessment, as the mask can distort the auditory speech signal.

2. 7. 2 (v) Rothenberg Mask

Ladefoged (1997) describes a pressure/flow system similar to Warren's PERCI system (1979). However, in the Ladefoged system the differential pressure is measured using a pharyngeal pressure tube instead of a nasal pressure tube. For this analysis, a small tube is passed through the nasal cavity so that it's open end rests on the back wall of the pharynx, 1 cm below the uvula. The other end of the tube is then attached to the pressure transducer. This system has been used in phonetic research, but would be too invasive for clinical use with young children. Ladefoged (1997) reports that variations in airflow are more difficult to measure than variation in air pressure using this system. He describes the use of the Rothenberg mask instead of the tube system for this purpose. This mask has a built in stainless steel gauze which allows for measurement of the rate of airflow by measuring the increase in pressure that occurs when air flows through the gauze. A divided Rothenberg mask permits the measurement of oral and nasal airflow simultaneously. However, the main disadvantage of the Rothenberg mask is that it is too large to provide an airtight seal between the oral and nasal cavities and around the face of children, under ten years.

2. 7. 3 Speech Stimuli

The speech sample used in testing is extremely important (Warren, 1979). The phoneme /p/ is used because it creates a stagnant column of air in the mouth and this eliminates the effects of tongue position or movement of air pressure and flow. It was recommended that the following speech sample be used to measure oral pressure, nasal pressure and oral flow: 'pa, pa, pa' repeated three times; 'pi, pi, pi' repeated three times and 'hamper, hamper, hamper' repeated three times (PERCI SARS Manual, 1994). Warren (1979) suggested that the nasal/plosive combination is helpful in determining whether a patient has a problem achieving closure rapidly. He also stated that the /mp/ combination in the word 'hamper' nearly approximates the degree of velopharyngeal closure, which occurs during fricative productions or continuous speech.

Zajac (1998, personal communication) reported no difference between pressure/flow measures on /pa/ and /pi/ for 152 normal speaking children. However, in a clinical population of adults and children, Smith and Guyette (1996) found that eight out of fifty one (15%) participants had velopharyngeal closing on /pa/ and excessive velopharyngeal openings on /pi/. This difference was explained by the authors by the fact that during the production of /i/ there is contraction of the palatoglossus muscle which may result in a downward movement of the palate. As there is no elevation of the tongue during the production of /a/, there is no contraction of the palatoglossus and no depression of the palate. In normal speakers, this movement of the palatoglossus is counteracted by the levator palatini. In individuals with borderline velopharyngeal function, the levator palatini cannot counteract the palatoglossus movement, possibly leading to velopharyngeal incompetence during /i/ (Moon, 1993). As a result of this finding, Smith and Guyette (1996) suggested that /pi/ be included in the speech sample for pressure/flow measures.

One of the main problems with the limited speech sample used in pressure/flow studies is that comparison with other assessment techniques is difficult. However, it has been argued that PERCI was not designed to evaluate speech performance. Rather, it provides information on differential pressures (Warren, 1979). Zajac, Mayo, Kataoka and Kuo (1996) adapted the technique to assess pressure/flow on /s/ in /si/. They positioned the catheter approximately 3 to 4 centimetres behind the central incisor with its opening perpendicular to the airflow. They attempted to achieve placement that was posterior to the anticipated point of lingual constriction at the alveolar ridge. They found that nasal pressure was unexpectedly greater than oral pressure during the production of /s/. This may have been due to the effect of the tongue obstructing the oral cavity. Zajac et al.

(1996) pointed out that placement of the catheter anterior to the point of lingual constriction would underestimate oral pressure relative to nasal pressure. This use of the pressure/flow system appears to be unreliable for the assessment of the /s/ phoneme.

2. 7. 4 Aerodynamic Studies Of Normal Speakers

Aerodynamic studies indicate a wide variation in pressure/flow measurements of normal speakers.

2. 7. 4 (i) Nasal Flow

Nasal airflow during the production of oral consonants has been found to range from 0 ml/s to 150 ml/s (Table 2.7). Previous studies have reported mean nasal airflow and the range of mean nasal airflow (Thompson & Hixon, 1979; Andreasson, Smith & Guyette, 1992; Zajac & Mayo, 1996).

Table 2.7 Nasal flow measurements for normal speakers during the production of different speech stimuli.

Study	**Speech Stimuli**	**Nasal Flow - Mean scores**	**Nasal Flow - Range of Individual scores**
Thompson & Hixon (1979)	/t/, /d/, /s/ & /z/	0 ml/s	0 - ±5 ml/s
Andreasson et al. (1992)	/p/ in 'pa'	6 ml/s (female) 17 ml/s (male)	0 - 42 ml/s 0 - 150 ml/s
	/p/ in 'hamper	12 ml/s (female) 30 ml/s (male)	0 - 42 ml/s 0 - 57 ml/s
	/m/ in 'hamper'	211 ml/s(female) 287 ml/s (male)	77 - 315 ml/s 169 - 442 ml/s
Zajac & Mayo (1996)	/p/ in 'hamper'	10 ml/s (female) 20 ml/s (male)	1 - 30 ml/s 2 - 76 ml/s
	/m/ in 'hamper'	129 ml/s (female) 149 ml/s (male)	69 - 216 ml/s 47 - 280 ml/s

Thompson and Hixon (1979) reported that the results indicated that on oral utterances (i.e. isolated utterance /i/ /s/ /z/, syllable repetitions /ti/ /di/ /si/ and /zi/ and the carrier phrase embedded syllables /iti/ /idi/ /isi/ and /izi/), nasal airflow was 0±5 ml/s for all participants. Andreasson et al. (1992) reported that mean nasal flow for one participant ranged from 0 to 150 ml/sec during the production of /p/ in 'pa'. This supports the conclusions of Laine et al. (1988) who suggested that 125 ml/sec be considered a cut-off value of nasal flow for the production of /p/ in 'hamper'. Interestingly, the participant in the Andreasson et al. study had nasal flow of 150 ml/sec during the production of /p/ in 'pa'.

Results of nasal flow measures for the production of /m/ in the study by Andreasson et al. (1992) differ from the results presented by Zajac and Mayo (1996). Zajac and Mayo (1996) stated that these differences were difficult to explain, however they pointed out that the differences may have been due to population size and procedural issues. This point will be discussed below in the section on gender influence.

Airflow rates in normal speakers were studied by Van Hattum and Worth (1967). They used a warm-wire flow meter with a divided mask which separated oral and nasal airflows. Measurements were difficult to make due to difficulty in locating consonants in syllables (Van Hattum & Worth, 1967). However, important facts regarding airflow measures were underscored in their study:

1. Voicing appeared to be the most important factor influencing expelled air volume (voiceless consonants displayed considerably more airflow than voiced consonants);
2. In comparing sounds in syllables and sounds in sentences, the expelled air volume tended to be less in sentence production;
3. Sounds at the beginning and end of the utterance were found to be more variable than sounds within the utterance;
4. Examination of percentage data (i.e. ratio of oral and nasal expelled air volume to total expelled air volume) indicated that on oral sounds a mean of 14% nasal expelled air volume was found and on nasal sounds a mean of 12% oral air volume was found.

Possible explanations for the presence of nasal air during the production of oral consonants were put forward by the authors. Firstly, there might have been a certain amount of anomalous airflow and the movement of the soft palate may have dispelled volumes of air which would have been measured as nasal airflow. Secondly, the movement of the closed lips inside the mask might also have created a similar displacement generated airflow or the mask may have had leakage of air. Thirdly, airflow did occur and closure of the nasal cavity during production of normal non-nasal vowels was sometimes incomplete. Finally, there may have been errors with the instrumentation (i.e. poor baseline and slow reaction time may also have been responsible for this finding). These findings underscored the problems with earlier approaches to aerodynamic assessment of nasal airflow. However, despite the errors in the study, important information regarding nasal airflow measurements was obtained for future studies. Measurements in the later pressure/flow measurements were never made on initial sounds, and voiceless plosives were used as speech stimuli

2. 7. 4 (ii) Pressure Measurements for Normal Speakers

Oral pressure for normal speakers has been reported in the literature (Table 2.8). Results indicated that normal intra-oral air pressure ranged from 2 to 20 cm H_2O during the production of /p/ in syllables 'pa', 'pi' and in the word 'hamper'. Pressure for /m/ ranged from 0 to 2.2 cm H_2O during production of syllables 'mi' and in the word 'hamper'.

Table 2.8 Range of mean oral pressure measurements for normal speakers during the production of /p/ and /m/.

Study	Speech Stimuli	Range of Mean
Zajac & Mayo, 1996	/p/ /m/	2.5 - 9.1 cm H_2O 0.1 - 4.1 cm H_2O
Andreasson et al., 1992	/p/ /m/	2 - 20 cm H_2O 0 - 2.2 cm H_2O

Differences between the results of the two studies will be discussed in the section on influence of gender on pressure/flow measures below.

2. 7. 4 (iii) Velopharyngeal Port Area Measurements for Normal Speakers

Velopharyngeal port area measurements were calculated from results of normal speakers using the Warren and DuBois equation of pressure difference between the nose and mouth and airflow through the nose. Warren (1967) initially described adequate velopharyngeal port area as below 0.05 cm^2. Subsequent studies on normal speakers reported smaller normal velopharyngeal port area measurements (Table 2.9).

Table 2.9 Range of velopharyngeal area port (vpa) measurements for normal speakers during the production of /p/ and /m/.

Study	Speech Stimuli	VPA Range
Andreasson et al., 1992	/p/ /m/	0 - 0.006 cm^2 0.01- 0.095 cm^2
Zajac & Mayo, 1996	/p/ /m/	0 - 0.002 cm^2 0.04 - 0.42 cm^2

2. 7. 4 (iv) Influence of Age on Pressure/Flow Measures in Normal Speakers

All the normative data reviewed provided information regarding pressure/flow measures for normal speakers including adults and children. However, in order to use this system with children, normal scores exclusively for children need to be established. Only one study provided data exclusively for normal speaking children. Zajac, Mayo and Kataoka (1997) assessed the developmental aspects of velopharyngeal function using pressure/flow measures. One hundred and fifty two children (aged 6 to 12 years) and 42 adults with normal speech were assessed using PERCI SARS. The following speech samples were used ‘pa’, ‘pi’, ‘mi’, ‘hamper’ and ‘peep into the hamper’. No difference in pressure/flow measurements for ‘pa’ and ‘pi’ were found. Mean oral pressure, nasal flow and velopharyngeal port area were calculated for four different age groups: 6 - 8 years; 9 - 10 years; 11 -12 years and adults (Table 2.10). Unfortunately, Zajac et al. (1997) did not report the range of values found in this population of normal speakers.

Table 2.10 Mean oral air pressure, nasal flow rates and velopharyngeal port area for normal speaking children (Zajac et al.,1997).

	Age group	Oral Pressure (cm H_2O) Mean (SD)	Nasal Flow (ml/s) Mean (SD)	VPA (mm^2) Mean (SD)
/p/ in 'pi'	6 - 8	8 (2.2)	1 (1)	0 (0.1)
	9 - 10	7.9 (2)	1 (1)	0 (0.1)
	11 - 12	7.2 (2.2)	2 (2)	0.1 (0.1)
	Adult	6 (1.4)	4 (7)	0.2 (0.4)
/p/ in 'hamper'	6 - 8	7.1 (2.5)	2 (3)	0.1 (0.2)
	9 - 10	7.2 (2)	5 (7)	0.2 (0.3)
	11 - 12	6.6 (1.2)	4 (3)	0.2 (0.1)
	Adult	5.3 (1.5)	15 (16)	0.7 (0.8)
/m/ in 'hamper	6 - 8	1.5 (0.7)	87 (38)	15.9 (10.7)
	9 - 10	1.5 (0.8)	96 (40)	15.3 (9.4)
	11 - 12	1.6 (0.8)	118 (47)	18.0 (10.8)
	Adult	1.3 (0.7)	139 (58)	19.8 (10.2)

Zajac et al. (1997) reported that for normal speakers, children use greater intra-oral pressure than adults and achieve greater velopharyngeal closure than adults. They also found that adults had greater mean nasal airflow during the production of /p/ in hamper than children. In a clinical study Dalston, Warren, Morr and Smith (1988) also found that children had higher oral pressure measurements than adults. Warren, Dalston and Mayo (1994) used a cut-off age of 15 years between children and adults. They reported that nasal cross-sectional size does not increase much after 15 years and that nasal cross-sectional area may influence resonance balance in pressure/flow measures. In their study they found that children in the borderline-inadequate and inadequate categories of velopharyngeal function were rated less hypernasal than their adult counterparts.

2. 7. 4 (v) Influence of Gender on Pressure/Flow Measures in Normal Speakers

Andreasson et al. (1992) found significant differences between scores for 10 males and 10 females in the following areas:

1. nasal flow measures for /p/ in 'hamper' ($t= 2.85$, $p = .01$)
2. velopharyngeal port area for /p/ in 'hamper'($t= 2.37$, $p = .03$)
3. nasal flow measures for /m/ in 'hamper' ($t = 2.46$, $p < .02$).

Male participants had significantly greater mean nasal flow measurements than female participants during the production of /p/ amd /m/ in 'hamper'. Males also had greater mean velopharyngeal port areas than females during the production of the word 'hamper'(Figure 2.7). No other differences between measures for male and females were significant. It should be noted that differences between male and female scores were only significant for the phonemes in the word 'hamper'. This highlights the importance of the phonetic context.

In contrast, Zajac and Mayo (1996) reported relatively small differences between scores for male and female speakers. The only measures that differed significantly were intra-oral pressure measures for /p/ ($t = -2.88$, $p = .007$). It may be that, although both studies found significant differences between male and female scores for some measures, the differences were marginal. Another explanation may be the wide variation in normal scores found in both studies. Zajac and Mayo stated that the differences in results may have been due to smaller number of participants in the study by Andreasson et al. (1992). The methodology was also different. In the study by Andreasson et al. (1992), the participants were asked to produce the speech sample 'hamper' three times in two or three groups. This instruction was ambiguous according to Zajac and Mayo (1996). In the study by Zajac and Mayo (1996), participants were asked to produce the word 'hamper' five times on one expiration. Zajac and Mayo (1996) speculated that the differences in respiratory and/or laryngeal function may have contributed to the different results from the two studies. Further studies on normal adults using a larger population may provide more conclusive results on gender differences.

2. 7. 4 (vi) Summary and Implications from Normal Aerodynamic Studies

Normal studies indicate a wide range of pressure/flow measurements for normal speakers. This review of the literature indicated that there are differences between pressure/flow measures for adults and children. Most of the normative data has been obtained from adults. In the present study it will be necessary to use normative data from children provided by Zajac (1998, personal communication). There are no reports in the literature of comparison between normal scores for children and scores from a clinical population of children.

2. 7. 5 Reliability of Pressure/Flow Measurements

Warren and DuBois (1964) reported that reproducibility of the pressure/flow data for normal participants based on sentence repetition was adequate. They stated that patterns of pressure and flow measurements were similar for the sentence repetitions. Few studies have evaluated test-retest reliability of the pressure/flow measurements.

Reliability of measurements was assessed by Zajac and Mayo (1996). When measurements of 2 participants with normal speech were repeated by the same scorer, a correlation coefficient of 0.99 was found. Reliability was assessed by comparing measurements obtained by one scorer with another. A correlation coefficient of 0.99 was reported. This study indicated that using the pressure/flow measurements, reliability of scores obtained was good. Unfortunately, no study assessed the reproducibility of the system using test-retest assessments.

2. 7. 6 Comparison of Pressure/Flow Measurements with Other Measurements

Results of pressure/flow measures have been compared with results of other techniques used in the assessment of speech problems related to velopharyngeal dysfunction. Fujiwara, Hiramoto and Kawano (1993) compared pneumotachography and videonasendoscopy in the assessment of velopharyngeal function. They used the aerodynamic technique described by Honjow et al. (1968) to assess oral pressure and nasal airflow rates. They found that the results of the combined pressure and flow measures related well to results of the videonasendoscopic assessment. This study indicated a good relationship between the direct assessment of velopharyngeal function (videonasendoscopy) and the indirect assessment of velopharyngeal function (pressure/flow measures) as described by Shprintzen (1995).

Validity of pressure/flow measurements has been assessed by comparing velopharyngeal port area, nasal flow and differential pressure measurements, with perceptual measures and nasalance scores. Warren (1967) analysed correlations between nasal airflow and velopharyngeal orifice size measures using his pressure/flow system. General correlation between measurements of nasal airflow and calculations of velopharyngeal orifice size was found to be 0.77. Laine et al. (1988) also reported good correlations between the two measures (Table 2.11). Although both studies reported good correlations for the entire study group, correlations varied when the study group was divided into categories according to velopharyngeal function. Results indicated weak correlations between nasal flow and velopharyngeal port area for borderline and inadequate groups.

Table 2.11 Relationship between velopharyngeal port area measurements and nasal flow measures

Study	Correlation Coefficient
Warren 1967	0.77
Laine et al., 1988	0.94

A small number of studies have assessed the relationship between pressure/flow measures and perceptual judgements of nasality using correlation analysis (Table 2.12). Results indicated varying relationships ranging from good (Dalston & Warren, 1986) to moderate (Warren et al., 1994) Discrepancies in correlations between perceptual judgements of nasality and velopharyngeal port area measurements may be explained by the different methodologies used in the two studies. Dalston and Warren (1986) compared a 5 point scale of perceptual assessment with three categories of velopharyngeal area measurements. In the Warren et al. (1994) study, comparisons were made between a 6 point perceptual scale and actual velopharyngeal area measurements. The wider range of scores for the velopharyngeal area may have reduced the level of correlation.

Table 2.12 Relationship between pressure/flow measures and perceptual judgements of nasality.

Study	VPA	Nasal Flow
Dalston & Warren, 1986	0.80	
Warren et al., 1994	0.66	0.61

Warren et al., (1994) found that the relationship between hypernasality and nasal airflow was highly significant for adults ($p < .001$); however, it was not significant for children ($p = .089$). They explained the differences in significance levels by the fact that children have smaller nasal airways. It was possible that high resistance diminished the need for increased respiratory effort which is generally required when the velopharyngeal mechanism is incompetent. Hence, the children may not have had increased nasal flow measurements. Another explanation may be that the airflow rates determined to represent mild and moderate hypernasality may be different for children. Warren et al. (1994) suggested that airflow rates of 180 ml/s were associated with ratings of mild to moderate hypernasality, while airflow rates of 300 were associated with moderate hypernasality. Much of the normal data reported in the literature prior to this study was obtained from adults.

Two studies assessed the relationship between pressure/flow measures and nasalance measures (Table 2.13). The results presented here indicate a large variation between estimates of the relationship between pressure/flow measures and nasalance scores. As indicated above, the difference in correlation may be explained by methodological differences.

Table 2.13 Relationship between pressure/flow measures and nasalance scores.

Study	Correlation Coefficient
Dalston & Warren, 1986	0.74
Dalston et al., 1991b	0.32

Further analysis of relationships between pressure/flow measures and other measurements will be discussed below in the section on sensitivity and specificity. First it is necessary to review cut-off values for the various pressure/flow measures presented in the literature.

2. 7. 7 Cut-off Values for Pressure/Flow Measurements

2. 7. 7 (i) Nasal Airflow Measurements

Warren (1967) reported that all participants classified as having velopharyngeal incompetency exhibited peak nasal airflow rates during the production of /p/ greater than 175 ml/s. All but two of the participants classified as having adequate closure produced an acceptable /p/ with less than 155 ml/s peak nasal flows. The remaining two had flow rates of 249 ml/s and 230 ml/s. Warren concluded that peak nasal flow rates above 250 ml/s during plosive sound production were considered to be indicative of inadequate closure. However, he specifies that the converse may not always be true. Flow rates of less than 250 ml/s may occur in spite of sphincter incompetency if there is nasal blockage or decreased respiratory effort.

Laine et al. (1988) reported that a nasal flow rate above 125 ml/s appears to be a fairly adequate predictor of velopharyngeal dysfunction. However, their study did indicate a high degree of variance. This cut-off value would be supported by normal studies (Andreasson et al., 1992; Zajac & Mayo, 1996), where the highest normal nasal flow rate was 150 ml/s.

2. 7. 7 (ii) Oral Pressure Measurements

Normal pressure ranges from 3 to 7 cm H_2O during the production of /p/ (Warren, 1979). Dalston et al. (1988) found that when participants were categorised according to velopharyngeal function (adequate, borderline and inadequate), pressure levels fell as the degree of inadequacy increased. However, even in the grossly inadequate group, the average oral pressure was 3.0 cm H_2O. Eighty seven percent of all participants assessed achieved intraoral pressure $\geq$ 3 cm H_2O. They compared oral pressure measurements in human speakers with velopharyngeal inadequacy and measurements from a vocal tract model with varying degrees of simulated velopharyngeal port openings. The human speakers were able to maintain oral pressure above 3 cm H_2O, whereas the model did not. This indicated that speakers may compensate for velopharyngeal inadequacy in order to maintain oral pressure. Warren (1986) suggested that pressure maintenance may be the primary goal of the vocal tract system. He stated that pressure maintenance may be achieved by changes within the vocal tract, such as increased respiratory effort, nasal grimace, high lingual carriage and increased glottal resistence.

2. 7. 7 (iii) Differential Pressure Measurements

The above results indicate that oral pressure measurements are not reliable for identifying participants with velopharyngeal inadequacy. As a result, differential pressure measurements have been used. Warren (1979) used a model of the vocal tract to simulate velopharyngeal opening. He estimated that differential pressure (i.e. difference between oral and nasal pressure) could be used to identify categories of velopharyngeal adequacy (Table 2.14).

Table 2.14 Summary of differential pressure measurements and adequacy of velopharyngeal function.

adequate closure	≥ 3 cm H_2O
borderline closure	1 to 2.9 cm H_2O
inadequate closure	< 1 cm H_2O

He studied 75 participants with varying levels of palatal competency and the data from these participants confirmed the above groupings. These groupings were confirmed in a later study by Morr et al. (1989). They reported that differential pressure measurements were fairly accurate predictors of adequacy or inadequacy of velopharyngeal function. But they concluded that the differential pressure score was only a screening index.

From the above studies it is evident that differential pressure measurements should provide more valid and reliable measures of airflow than oral or nasal pressure measurements on their own.

2. 7. 7 (iv) Velopharyngeal Port Area Measurements

Warren (1979) reported that when openings of the velopharyngeal port were less than 0.05 cm^2, voice quality was within normal limits. He found that openings between 0.05 and 0.1 cm^2 usually did not interfere with production of intraoral pressure on plosives; however, some nasal emission could occur. Openings between 0.1 and 0.19 cm^2 represented borderline velopharyngeal function where some speakers sounded normal while others had slight to moderate nasal emission and hypernasality (Table 2.15). Similar cut-off values for velopharyngeal incompetency were identified by Isshiki, Honjow and Morimoto (1968). Results indicated that the critical size of velopharyngeal closure necessary for acceptable speech was 5 mm in diameter. However, they pointed out that there was no definite point of the velopharyngeal dimension where the speech suddenly changed from normal to abnormal. It was evident that the degree of nasality depended not only on the size of the velopharyngeal sphincter but also on several other factors such as mouth opening and position of the tongue.

Table 2.15 Summary of velopharyngeal (VP) function related to the size of the velopharyngeal port area (VPA) (Warren, 1967; Warren, 1979).

VP Function	VPA
adequate closure	< 0.05 cm^2
adequate-borderline	0.05 cm^2 to 0.09 cm^2
borderline - inadequate	0.10 cm^2 to 0.19 cm^2
slightlyinadequate closure	0.20 cm^2 to 0.40 cm^2
moderate inadequacy	0.41 cm^2- 1 cm^2
gross inadequacy	> 1 cm^2

Results of studies which identified cut-off values to distinguish between normal and abnormal speakers and the results of normal studies indicate that there is large variation of individual scores for each measurement made. Therefore, this information should be used as a guideline. Most of the studies which identified cut-off values based these measurements from adult data or from model analogue data. Hence, these cut-off values may only be used as a guideline for children. It will be necessary to compare normal pressure/flow measurements with measures from children with speech problems in order to identify cut-off values or range of values for children.

2. 7. 8 Sensitivity and Specificity

Dalston et al. (1991b) compared nasalance scores and velopharyngeal port area as measured by the pressure/flow technique. They found that using a nasalance cut-off value of 32% and a velopharyngeal area cut-off of 0.10 cm^2 , the Nasometer had a test sensitivity of 0.78 and a specificity of 0.79. When perceptual judgements were compared to the Nasometer, test sensitivity and specificity were higher. However, this was not surprising as the Nasometer and the pressure/flow system measured different phenomena. One measured acoustic energy and the other measured aerodynamic performances of the velopharyngeal mechanism (Dalston & Warren, 1986). It would have been useful to compare nasalance scores with other pressure/flow measures such as nasal flow.

Comparing the different measurements obtained from the pressure/flow technique, sensitivity and specificity were found to be good. Laine et al. (1988) assessed the screening of velopharyngeal closure based on nasal airflow rate measurements. When groups were categorised according to adequate and inadequate categories, sensitivity was 0.85 and specificity was 0.96. Sensitivity and specificity of differential pressure measurements were reported as 0.88 and 0.94 respectively when compared to measurements of velopharyngeal port area (Morr et al., 1989). They found that, in general, when the differential pressure was above 3 cm H^2O the velopharyngeal port area was less than 0.01 cm^2. However the borderline-inadequate group demonstrated a substantial scatter of differential pressure values. One would have expected good agreement between velopharyngeal port area and other pressure/flow measures, as the differential pressure and nasal flow measurements were used to calculate velopharyngeal port area. However, verification of the categories of velopharyngeal port area could have been made with perceptual judgements.

Pressure/flow techniques were used to evaluate the results of pharyngeal flap surgery (Smith, Skef, Cohen & Dorf, 1985). Thirty one patients who had undergone pharyngeal flap surgery were selected and perceptual judgements of speech were also obtained following surgery. Results indicated that, using pressure/flow criteria, 52% of outcomes were considered successful and 35% were considered to have substantial nasopharyngeal airway obstruction. Perceptual judgements were based on an individual listener's judgement of oral/nasal balance during single word production, sentence production and conversational speech. In 85% of participants, nasality judgements were comparable to the nasal flow findings. In 10% of the participants, resonance quality was judged hyponasal while pressure/flow patterns indicated adequate valving. The authors suggest that discrepancies may be due to poor validity of nasality judgements, as only one judge was used in the study and the speech sample would be influenced by a phonetic environment and subject type. They conclude that perceptual and nasal flow data can provide important information in evaluating velopharyngeal function post-operatively.

Correlation studies in the literature evaluated the relationship between pressure/flow measures and measures of velopharyngeal function (Fujiwara et al., 1993), acoustic measures of nasality (Dalston et al., 1986) and perceptual ratings of nasality (Dalston et al., 1986; Warren et al., 1993a). No study reviewed evaluated the relationship between pressure/flow measures with perceptual ratings of nasal airflow. One would expect a good correlation between pressure/flow and perceptual ratings of nasal airflow as both techniques measure airflow, but in different ways.

2. 7. 9 Timing Studies

Under certain circumstances the perception of hypernasal speech may be related more to the amount of time the orifice is open than to the actual amount of air passing through the nose or to the extent of velopharyngeal opening (Warren, Dalston & Mayo, 1993a). Hypernasality may be a time related phenomenon and aerodynamic changes that occur as a result of velopharyngeal incompetency may not be linearly related to perceptual judgements of oral/nasal resonance. Warren et al. (1993a) reported that when duration of the velar opening and closing movements prior to and after a nasal consonant is too long, speech is often judged to be hypernasal. Conversely, if the opening and closing duration is too short, speech may be described as hyponasal. Hence, the amount of time the velopharyngeal sphincter is open or closed may have a more significant effect on resonance than the amount of airflow passing through the nasal cavity.

Recently the pressure/flow technique has been used to determine whether the temporal characteristics of aerodynamic data associated with utterances could differentiate speakers with palatal incompetence. Warren, Dalston, Trier and Holder (1989) analysed the /mp/ consonant blend in the word "hamper" for 70 participants - 10 normal speakers, 20 with adequate velopharyngeal closure, 20 with borderline velopharyngeal closure and 20 with inadequate velopharyngeal closure. Timing variables for the word 'hamper' included: (1) beginning of airflow, (2) peak airflow, (3) end of airflow, (4) beginning of pressure, (5) peak pressure and (6) end of pressure. Figure 2.7 illustrates the normal airflow-pressure relationship for the speech production of the word 'hamper". Results indicated that each of the four groups manifested unique timing characteristics in their pressure/flow tracings. Typical timing characteristics of the velopharyngeal dysfunction included an overlap between the airflow graph and the pressure graph (Figure 2.8). There was a shorter time gap between the beginning of the flow graph and the beginning of the pressure graph, the middle peak of the flow graph and the middle of the pressure graph and the end of each graph.

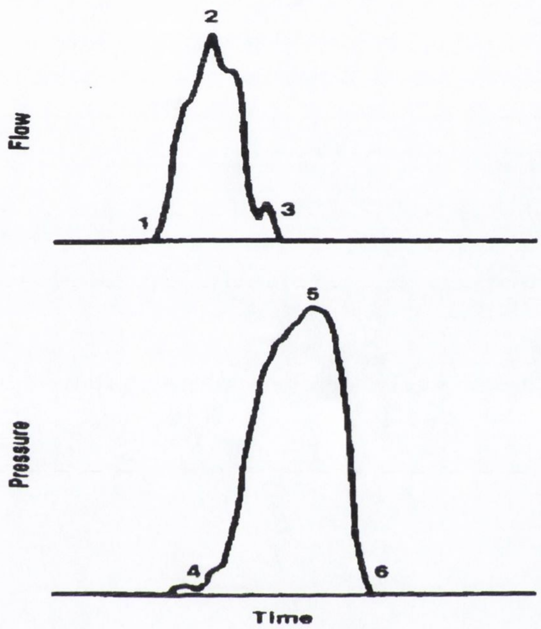


Figure 2.7 Normal graph indicating timing between flow and pressure graphs during the production of the /mp / combination in 'hamper'. (1) beginning of airflow, (2) peak airflow, (3) end of airflow, (4) beginning of pressure, (5) peak pressure and (6) end of pressure.

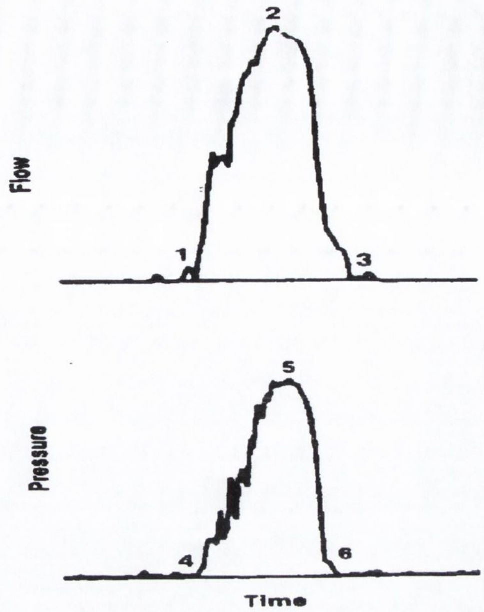


Figure 2.8 Pressure/flow graph of a speaker with velopharyngeal dysfunction during the production of /mp/ in 'hamper'. (1) beginning of airflow, (2)peak airflow, (3) end of airflow, (4) beginning of pressure, (5) peak pressure and (6) end of pressure. Note the overlap of the flow and pressure graphs.

In the same study, Warren et al. (1989) divided their participants into two groups according to perceptual ratings of nasality. They found that the group with normal or near normal ratings of nasality exhibited timing characteristics similar to normal and adequate groups. The group with mild/moderate to severe ratings of nasality had timing characteristics similar to the velopharyngeal dysfunctional group.

In a further study, Warren, Dalston and Mayo (1993b) reported on hypernasality in the presence of adequate velopharyngeal closure. This study indicated that there were unique timing features that differentiated the hypernasal but adequate group from the two control groups. These features included a delay of approximately 15 milliseconds in achieving velopharyngeal closure, a longer interval of nasal airflow and a shorter duration of velopharyngeal closure. In this study the authors noticed that participants with borderline velopharyngeal closure, who were lwss hypernasal had timing patterns similar to normal participants. Participants with greater hypernasality had timing patterns similar to patients with inadequate velopharyngeal closure. They found that particular timing variables were especially useful in discriminating between the three groups investigated.

It is apparent from the timing studies that the temporal characteristics of pressure/flow measures provide important clinical information regarding speakers with velopharyngeal dysfunction. Two studies, however, indicated problems with this approach. One study indicated that the temporal measures reported to date do not predict the clinical categorisation of velopharyngeal function (Dalston et al., 1991c). Another study reported that temporal measures had a weaker correlation with perceptual judgements ($r = 0.53$) than other pressure/flow measures (Warren et al., 1994). Although temporal characteristics may provide useful diagnostic information, in the future further investigation into this type of pressure/flow analysis is required. It is possible that in cases where perceptual ratings of nasality and/or nasal airflow are high, but nasal flow, differential pressures, or velopharyngeal port area measures are normal, the temporal characteristics may be more closely associated with perceptual ratings. This notion is supported by the study by Warren et al. (1993b).

2. 7. 10 Conclusions from Pressure/Flow Studies

Review of the literature indicates that pneumotachographic systems of airflow and pressure measures are more reliable and versatile than the simpler anemometry systems (Counihan, 1971; Warren, 1975). The pressure/flow system described by Warren and DuBois (1964) has been used extensively for research purposes during the past three decades in the US. This system assesses airflow and pressure during the production of the bilabial /p/ in isolated syllables and in the word "hamper". The PERCI SARS system can measure various pressures, flows, areas and resistances. It can be used to assess timing characteristics as well as vocal parameters in detail (Riski et al., 1995).

Warren (1979) found that the nasal/plosive combination in "hamper" allows for assessment of rapid closure of the velopharyngeal sphincter. Information from the study by Smith and Guyette (1996) indicated that the syllables /pa/ and /pi/ should be included in the speech sample as well as the word ‘hamper’, as they reported differences in a clinical population when the plosive /p/ was followed by a high vowel.

Cleft palate participants were divided into categories according to velopharyngeal function - adequate, slight inadequacy, moderate inadequacy and gross inadequacy in Warren's (1967) study These groupings have been confirmed in subsequent studies (Warren, 1979; Laine et al., 1988) and velopharyngeal gap size and pressures have been estimated for each group. Critical velopharyngeal gap size was found to be 5 mm in diameter (Isshiki et al., 1968) or 0 to 0.2 cm^2 (Warren, 1967). Differential pressure measures below 3 cm H_2O have been reported to indicate velopharyngeal inadequacy (Warren, 1979; Morr et al., 1989). Studies of normal speakers using the pressure/flow technique indicated a wide range of normal scores (Andreasson et al., 1992; Zajac & Mayo, 1996). Although mean values have been identified for adults (Warren, 1979; Laine et al., 1988; Andreasson et al., 1992; Zajac & Mayo, 1996) and children (Zajac & et al., 1997), no cut-off values to distinguish between normal and abnormal speaking children have been established.

Studies indicated a moderate to good relationship between pressure/flow measures and other measures of speech and velopharyngeal function (Dalston & Warren, 1986; Warren et al., 1994). None of the studies reviewed compared perceptual ratings of nasal airflow with pressure/flow measures.

Assessment of temporal characteristics of aerodynamic data indicates that duration of velopharyngeal opening prior to or after a nasal consonant may be related to perception of hypernasality (Warren et al., 1989; Warren et al., 1993a; Warren et al., 1993b). Warren et al. (1993a) concluded that the extent of time the nasal chamber is open may be more important than the actual amount of air escaping into the nose. There are still problems with temporal measures which need further investigation.

2. 8 General Conclusions

One of the basic problems identified in the literature on the assessments of nasality and nasal airflow problems has been the lack of definition of terms. This problem was identified both with regard to the terms used to describe nasality and nasal airflow in pathological speech, as well as with regards to the terms used in relation to reliability and validity of measurements (Bzoch, 1989; Cordes, 1994). Based on the literature and clinical experience, definitions have been presented in order to develop a descriptive perceptual profile for assessments of nasality and nasal airflow problems. The terms reliability, agreements, validity, cut-off values, test sensitivity and specificity have been defined in order to review the literature of assessment techniques, and to identify appropriate methods for evaluating the perceptual framework and instrumentation in the present study.

The central role of the perceptual assessment has been underscored. It is generally accepted that perceptual assessments of nasality and nasal airflow are the " final arbiter in the decision making" and the standard against which instrumental measures are evaluated (Kent, 1996; p. 7). However, perceptual assessments presented many problems in terms of reliability, agreements and validity. In his review of instrumental assessments of velopharyngeal valving, Sphrintzen (1995) points out that the search for new procedures is fuelled by the fact that each individual procedure currently in use has some type of drawback resulting in examiner dissatisfaction. This point also explains why instrumental assessments of nasality and nasal airflow have been developed in order to overcome some of the problems of perceptual assessments. One of the ways of overcoming the limitations of perceptual analysis of speech is to supplement perceptual ratings with instrumental analysis of the same behaviour (Kent, 1996).

Studies on the use of nasometry in the evaluation of nasality, have indicated that the Nasometer is a useful tool (Dalston et al., 1991b). Certain limitations have been identified. Because of language and dialectal variations in normal nasometry, nasalance scores for normal speakers in a population need to be established. Studies have used varying speech stimuli in nasometry with varying results. It is generally accepted that a passage devoid of nasal consonants is useful in identifying patients with hypernasality, and a passage containing nasal consonants is useful in identifying patients with hyponasality. The value of high pressure and low pressure consonant sentences has been raised as an issue. Varying results of the relationship between perceptual judgements of nasality and nasalance scores have been presented. However, in more recent studies, sensitivity and specificity of the Nasometer have provided encouraging results.

The review of literature indicated that the pressure/flow technique described by Warren (1967; 1979) is the most useful aerodynamic approach in assessing disorders of speech. One of the major drawbacks of the systems is the limited speech sample used during assessments. However, cut-off values for the various pressure/flow measurements have been found to identify participants with velopharyngeal inadequacy (Warren, 1979; Laine et al., 1988; Morr et al., 1989). Studies have identified a normal range of scores for normal speakers (Andreasson et al., 1992; Zajac & Mayo, 1996). Only one study has reported values exclusively for normal children (Zajac et al., 1997). Research on the relationship between pressure/flow measures and perceptual judgements is limited. No study has compared pressure/flow measurements with perceptual judgements of nasal airflow.

It is evident from the foregoing sections that there is no one ideal measurement of nasality and nasal airflow. Therefore the issue is not which assessment technique is the best (Moon, 1983). The underlying premise is that a combination of different assessment techniques would provide the most insights about any particular experimental work or clinical question (Folkins & Moon, 1990).

2. 9 Statement of the Problem

The review indicates the difficulties in the assessment of nasality and nasal airflow problems associated with cleft palate, non cleft velopharyngeal dysfunction and nasal obstruction. No one assessment tool will adequately assess these speech problems. The aim of the present study is to develop an assessment protocol which reliably assesses nasality and nasal airflow in children. In order to develop such a protocol the following areas require further investigation:

1. The reliability and validity of a descriptive scale of nasality and nasal airflow developed according to working definitions.

2. The relationship between nasalance values for normal Irish English-speaking children and nasalance values for a clinical population.

3. The relationship between pressure/flow measurements for normal children and pressure/flow measurements for a clinical population.

4. The relationship between perceptual ratings on the descriptive scale and nasalance scores.

5. The relationship between perceptual ratings on the descriptive scale and pressure/flow measurements.

6. The value of specific speech stimuli in identifying aspects of nasality and nasal airflow problems.

7. Cut-off values need to be established for the instrumentation in order to distinguish between normal and pathological speakers.

It is hoped that by combining the perceptual, acoustic and aerodynamic approaches to assessment, some of the problems of assessment of nasality and nasal airflow may be understood and even overcome.

CHAPTER 3

NASALANCE SCORES FOR NORMAL IRISH CHILDREN

Introduction

Instrumental techniques for the assessment of speech problems associated with velopharyngeal insufficiency have been developed during the past decade (Fletcher et al., 1989; Warren et al., 1993). One of these instruments, the Nasometer, has been used clinically and in research as a non-invasive method for assessing nasal resonance. This microcomputer-based instrument provides an objective measure of nasality by measuring oral and nasal acoustic energy during speech production and calculating a 'nasalance' score. Nasalance is a ratio of nasal acoustic energy to the sum of nasal plus oral acoustic energy multiplied by 100.

Since the Nasometer was introduced in 1986, several studies have reported the usefulness of the Nasometer in assessment of nasality problems associated with velopharyngeal insufficiency (Dalston et al., 1991b; Nellis et al., 1992) and with nasal obstruction (Dalston et al., 1991a). Other studies compared nasalance scores with various assessments of nasality and velopharyngeal function with conflicting results (Dalston, 1989; Williams et al., 1990; Nellis et al., 1992). These studies concluded that the Nasometer is a useful clinical tool for aiding assessment and diagnosis of nasality problems when used as a supplementary assessment tool.

Studies of normal speakers indicated that nasalance scores vary across languages (Haapanen, 1991a; Anderson, 1996; Trindade et al., 1997) and across dialects (Seaver et al., 1991). Seaver et al. (1991) found significant differences in nasalance scores for speakers of American English and suggested that clinicians may need to establish regional norms in order to assess patients with nasality problems. To date no normal nasalance scores have been reported for English speakers in the United Kingdom or Ireland. Previous studies have indicated controversy regarding gender differences in mean nasalance scores in normal speakers. Seaver et al. (1991) found small but significant differences in mean nasalance scores between male and female speakers, whereas as Trindade et al. (1997) and van Doorn and Purcell (1998) reported no differences in nasalance scores between male and female. No study evaluated gender difference in the nasalance scores specifically for children.

Several speech samples are included in the nasometry package for use in assessment of pathological speech. Previous research has shown that the Zoo Passage (devoid of nasal consonants) is useful in identifying individuals with velopharyngeal dysfunction (Dalston et al., 1991a), while the Nasal Passage is useful in identifying individuals with hynonasality (Dalston et al., 1991b). The Rainbow Passage (which contains 11% nasal consonants) is believed to represent the percentage of nasal consonants in conversational speech (Fletcher et al., 1989; Dalston & Seaver, 1992). However Dalston and Seaver (1992) reported poor correlation between nasalance scores on the Rainbow Passage and perceptual judgements of nasality. Another problem with the Rainbow Passage was that is it has been found to be too difficult semantically and syntactically for young children (Watterson et al., 1993). Waterson et al. (1993) also reported that the Zoo Passage was difficult for children to repeat. Some simplified passages have been developed for use with the Nasometer (MacKay & Kummer, 1994). However, a pilot study of normal speaking Irish children indicated that these passages were also difficult to repeat due to cultural differences (Appendix 6). A speech sample which included separate High Pressure consonant sentences and Low Pressure consonant sentences has been recommended by Karnell (1995). He stated that, when nasal turbulence was present, nasalance scores on High Pressure consonant sentences may be artificially high. The elevation of nasalance scores on High Pressure consonant sentences may become apparent in children with nasal emission and/or nasal turbulence, if separate nasalance scores are obtained for High Pressure consonant and Low Pressure consonant sentences.

The aim of the present study is to obtain normal nasalance scores for English-speaking Irish children, using a speech sample that contains all consonant types, with 11% nasal consonants, representative of normal conversational speech (Dalston & Seaver, 1992). These data will be used as a baseline against which nasalance scores of children with abnormal nasality can be compared. The speech sample should also be analysed in categories according to sentence type, so that separate measurements for High Pressure consonant sentences and Low Pressure consonant sentences are made. The speech samples should be sufficiently easy for young children to repeat and should be without cultural bias. The study also aims to evaluate gender differences in mean nasalance scores in children.

A pilot study was carried out to ascertain which speech samples should be used. The sample consisted of sixteen test sentences adapted from the 'GOS.SP.ASS: A Screening Assessment of Cleft Palate Speech' (Sell et al., 1994) which is designed to be used in the clinical assessment of children with nasality and nasal airflow problems. In the pilot study, 26 children with normal speech were assessed on the Nasometer model 6200.3 repeating the 16 test sentences. Results indicated that the sentence sample was useful in obtaining normal scores. However, the order and timing of presentation had to be changed to optimise the results obtained (Appendix 6. Nasometry Pilot Study).

3. 1 Methodology

Participants

The teachers from eight different classes in a normal national school were asked to select ten children who they considered to have normal speech and hearing. An experienced Speech and Language Therapist listened to each child repeat the test sentences. Seventy children (36 girls and 34 boys, aged 4;11 years to 13 years) were found to have no articulation errors and normal resonance. All children had an Irish-English accent. None of the children had a history of cleft palate, velopharyngeal dysfunction, nasal obstruction, hearing problems, neurological involvement, significant medical problems or speech therapy.

Instrumentation

The Nasometer (6200.3) is manufactured by Kay Elemetrics. It consists of a headset, containing a sound separator with microphones on either side, which detects oral and nasal components of the participants' speech. The sound signal is filtered and digitised and the data is processed by a computer (IBM PC). The resultant signal is the ratio of nasal to nasal-plus-oral acoustic energy and is expressed as a Nasalance score.

Speech sample

The speech sample was adapted from the original Great Ormond Street Speech Assessment (Sell et al., 1994). The sixteen test sentences (Total Test Sentences) included all consonant types and had 11% nasal consonants. The sample also included a group of sentences which contained high pressure consonants (plosives, fricatives and affricates) and was devoid of nasal consonants, representing a speech sample which is vulnerable to the effects of velopharyngeal dysfunction (D'Antonio & Scherer, 1995).

Sixteen test sentences were recorded for all children. The sentences were recorded in the following order to allow for detailed analysis according to sentence category:

i. High pressure consonants (devoid of nasal consonants)
ii. Low pressure consonants (devoid of nasal consonants)
iii. Mixed consonants (nasal and oral consonants)
iv. Nasal consonants (55% nasal consonants)

(Appendix 7)

Experimental Procedure

The data collection took place over two days. The Nasometer was calibrated at the beginning of each day and calibration was checked after the testing of each group of 15 participants. The procedure was carried out in a small quiet room away from the main classroom. The headset was placed on each child's head, with the sound separator below the child's nose and above the upper lip. The child repeated each of the 16 test sentences. The sentences were presented and recorded in groups according to sentence category. Recordings were made at 4 second time display, with a 1 second space between each sentence within the category and a 5 second space between each category to allow for subsequent identification of categories for analysis.

The mean nasalance score for the 16 test sentences was calculated using the Nasometer software. Subsequently, each category was marked on the visual display on the computer terminal by placing a cursor at the beginning and at the end of the category. The mean nasalance score was then calculated for the marked area providing a nasalance score for each sentence category. A repeated measures two-way analysis of variance was carried out to evaluate the interaction between sentence categories and gender. A Bonferroni post-hoc analysis was carried out to ascertain if there was significant difference between sentence categories.

3. 2 Reliability

Ten participants were randomly selected for test-retest analysis. The 10 participants were asked to repeat the 16 Total Test sentences immediately after initial testing. The headset was not removed, but was checked to ensure appropriate placement. Reliability coefficients were calculated for each sentence category, using Generalizability Theory (Streiner & Norman, 1995). Results indicated good reliability for the Total Test sentences, High Pressure consonant sentences and Low Pressure consonant sentences. Reliability for the Mixed sentences and the Nasal sentence was weak (Table 3.1). Scattergrams of individual variations for each sentence category are presented in Appendix 8.

Table 3.1 Test-retest reliability for nasalance scores on all sentence categories.

	Total Test	**High Pressure**	**Low Pressure**	**Mixed consonant**	**Nasal**
Reliability Coefficient	0.73	0.83	0.74	0.52	0.15

Inspection of the data indicated that 100% of mean nasalance scores for the Total Test sentences were within 4 percentage points of the score on repetition of the sentences. Detailed analysis indicated that for repetition of high pressure consonants, low pressure consonants and mixed consonants, there was individual variation of up to 5 percentage points. Ninety percent of mean nasalance scores were within 5 percentage points of the score during repetition of the Nasal sentence. These results are similar to those reported by Seaver et al. (1991) and van Doorn and Purcell (1998). When the one participant who had a large variation in test-retest scores was removed from the group, the reliability coefficient for the Nasal sentence improved to 0.79.

3. 3 Results

3. 3. 1 Results of Total Sentence Sample

The mean nasalance score for the group (n = 70) during the production of the 16 Total Test sentences was 26.6%, with a standard deviation of 5% and ranged from 17% to 35% (Figure 3.1).

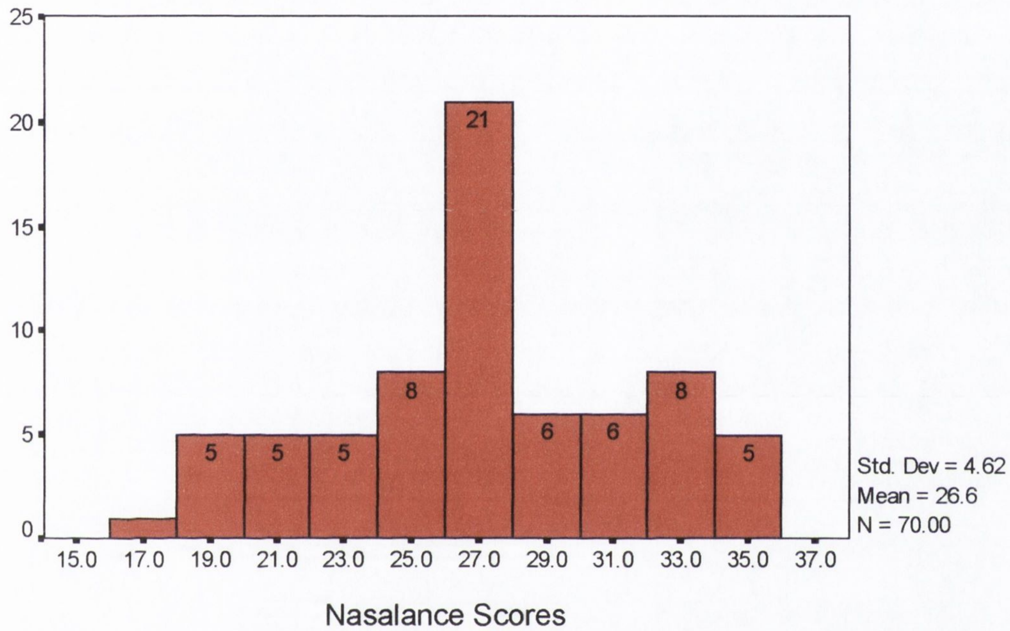


Figure 3.1 Nasalance scores for normal speakers during production of the Total Test sentences.

3. 3. 2 Results of Separate Speech Categories

Detailed analyses of the 16 sentences were carried out to obtain nasalance scores for each sentence category. The mean, standard deviation and range of nasalance scores for the group during production of each sentence category was calculated (Table 3.2).

Table 3.2 Mean nasalance scores, standard deviations and range of nasalance scores for each sentence category for 70 normal Irish English-speaking children.

Sentence Category	**Mean**	**SD**	**Range**
High Pressure consonant sentences	14%	5%	7% - 25%
Low Pressure consonant sentences	16%	6%	7% - 30%
Mixed consonant sentences	34%	7%	19% - 47%
Nasal sentence	51%	7%	33% - 68%

Similar means and range of nasalance scores were found for High Pressure consonant sentences and Low Pressure consonant sentences which were devoid of nasal consonants. There were higher means and a greater range of nasalance scores for the Mixed consonant sentences and the Nasal sentence (Figure 3.2).

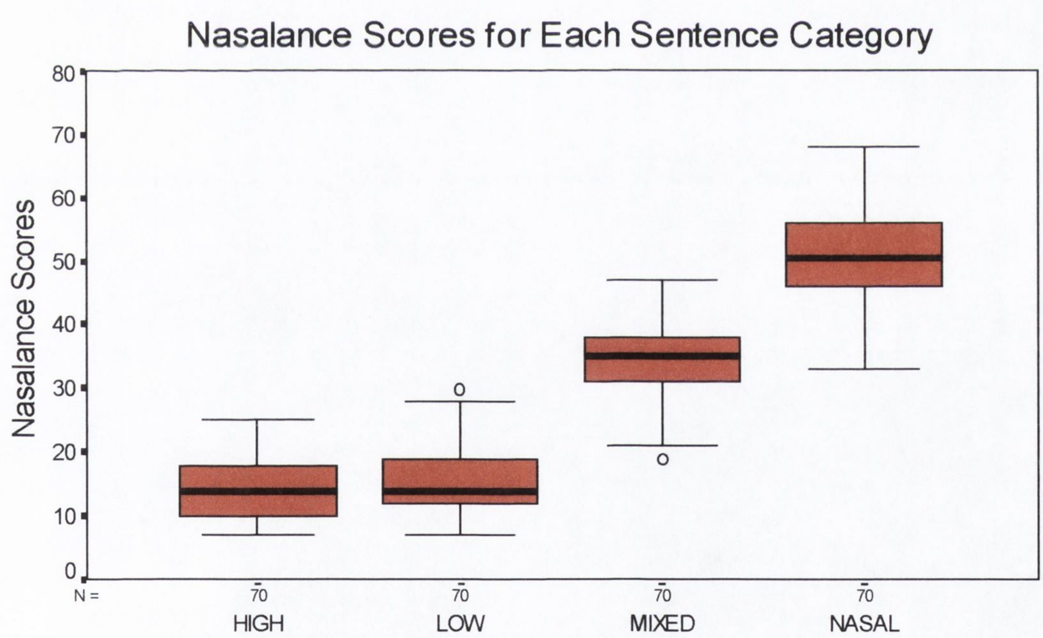


Figure 3.2. Boxplot indicating the median and range of nasalance scores for the four sentence categories. The graph indicates the median nasalance score (black line), 50% of participants who fall within the 25th and 75th percentile (red box), the range of scores (thin line) and outliers (o) for each sentence category - high (High Pressure sentence category), low (Low Pressure consonant sentence category), mixed (Mixed consonant sentence category) and nasal (Nasal sentence).

A repeated measure two-way analysis of variance indicated no significant interaction between sentence category and gender (f = 1.54, df = 4,272, p = 0.1). There was no significant difference in nasalance scores between male and female speakers (f = 0.55, df = 1,272, p = 0.46). However, there was a significant difference in nasalance scores for the different sentence categories (f = 1271, df = 4,272, p ≤ .001) (ANOVA Table Appendix 9). A Bonferroni post-hoc analysis indicated significant differences between each sentence category (p ≤ .001), except between High Pressure consonant sentences and Low Pressure consonant sentences (p = .09) (Table 3.3).

Table 3.3 Bonferroni post-hoc analysis of difference between sentence categories indicating differences, standard error, p value and confidence intervals.

Sentence Type	**Difference**	**Standard Error**	**p value**	**Confidence Interval**
Total - High	-12.21	.5969	≤.001	10.42 - 13.89
Total - Low	-10.67	.5969	≤.001	8.98 - 12.36
Total - Mixed	7.54	.5969	≤.001	5.85 - 9.23
Total - Nasal	24.51	.5969	≤.001	22.82 - 26.20
High - Low	1.54	.5969	= .09	-0.14 - 3.24
High - Mixed	19.75	.5969	≤.001	18.06 - 21.44
High - Nasal	36.72	.5969	≤.001	35.03 - 38.41
Low - Mixed	18.21	.5969	≤.001	16.52 - 19.90
Low - Nasal	35.18	.5969	≤.001	33.49 - 36.87
Mixed - Nasal	16.97	.5969	≤.001	15.28 - 18.66

3. 4 Discussion

Studies of gender differences in nasalance scores have been inconclusive in adults, although perhaps more conclusive for children. Seaver et al. (1991) report significantly higher nasalance scores among normal female adults than normal male adults. However, Litzaw and Dalston (1992) found no gender difference. The latter study had a smaller normal sample of 30, whereas Seaver et al. (1991) had a sample of 148. The present study examined the gender differences in children and found no difference in nasalance scores between girls and boys. This finding supports previous studies on children where no difference was found (Watterson et al., 1996; Trindade et al., 1997; van Doorn & Purcell, 1998).

Mean nasalance scores for speech samples containing all consonants and 11% nasal consonants (i.e., a sample representative of conversational speech), are reported in the literature (Fletcher et al., 1989; Seaver et al., 1991; Watterson et al., 1996) (Table 3.4). The mean nasalance for American normal adults and children on the Rainbow Passage was found to be 36% in two different studies (Fletcher et al., 1989; Seaver et al., 1991), whereas the mean nasalance score in this study was 26%. Therefore, the Irish and American mean nasalance scores differ by 10% when the test sample contains an equivalent percentage of nasal consonants and is representative of conversational speech. Such difference is not observed in speech samples devoid of nasal consonants (see below). This provides some evidence for the observation made by Wise (1957). He reported that English Phoneticians described the American dialects as nasal, unlike American Phoneticians. The present comparison of normal nasalance scores in the speech sample that may be considered representative of conversational speech supports the English Phoneticians' perceptual judgements of a higher degree of nasal resonance in American speakers compared to Irish speakers. (Rainbow, Zoo and Nasal Passages are presented as assessment passages with the Nasometer and referred to in other studies - Appendix 3).

Standard deviation in the present study was 5% which compares well with the American studies. This would imply that, although American speakers have higher nasalance than Irish-English speakers, the variation around the mean is similar.

Table 3.4. Nasalance norms for passages containing all consonant types. The table indicates the mean nasalance scores and standard deviations for speakers during the production of "similar" speech samples. The number of participants in each study is illustrated. Results from the present study are included in italics for comparison.

Participants	**Investigator**	**N**	**Speech Sample**	**Mean nm**	**SD**
American Children	Fletcher et al., 1989	117	Rainbow	36%	5%
American Adults	Seaver et al., 1991	148	Rainbow	36%	6%
American Children	Watterson et al., 1996	20	Mouse	32%	7%
Irish Children	*Sweeney et al.*	*70*	*16 sentences*	*26%*	*5%*

Results indicated that there was a significant difference between mean nasalance scores across all sentence categories. A highly significant difference was found between the Total Test sentences and each individual sentence category. This can be explained by the different percentages of nasal consonants in each sentence category. Analysis indicated that there was no significant difference between normal nasalance scores for High Pressure consonant sentences and Low Pressure consonant sentences. The lack of difference may be explained by the absence of nasal consonants in the two categories. As expected, the test sentences containing nasal consonants had significantly higher nasalance scores than the test sentences devoid of nasals (i.e. High Pressure and Low Pressure consonant sentences). The range of scores for the group also increased when the sentences contained nasal consonants.

The mean nasalance scores (14%) for High Pressure consonant sentences compares well with previous data on speech samples containing no nasal consonants and approximately 75-80% high pressure consonants (Table 3.5). In previous studies of English-speakers the mean nasalance score was between 13% and 16% (Fletcher et al., 1989; van Doorn & Purcell, 1998). The present study showed little difference between Irish English-speakers and American speakers when the speech sample was devoid of nasal consonants.

Table 3.5. Nasalance norms for sentences containing high pressure consonants and devoid of nasal consonants comparable to the Zoo Passage. The number of participants in each study is illustrated. Results from the present study are included in italics for comparison.

Participants	**Investigator**	**N**	**Speech Sample**	**Mean NM**	**SD**
American Children	Fletcher et al., 1989	117	Zoo	16%	5%
American Adults	Seaver et al., 1991	148	Zoo	16%	7%
Finnish Adults/ Children	Haapanen, 1991a	42	1 Sentence	14%	5%
American Children	Watterson et al., 1996	20	Turtle	16%	3%
Brazilian Portuguese Children	Trindade et al., 1997	20	Brazilian Portuguese Zoo	9%	3%
Australian Children	van Doorn & Purcell, 1998	245	Zoo	13%	6%
Irish Children	*Sweeney et al.*	*70*	*5 sentences*	*14%*	*5%*

Similar nasalance scores were found for Irish English-speakers and American English speakers on passages devoid of nasal consonants. The presence of nasal consonants in the speech sample appears to have a greater effect on the overall resonance of American speakers as compared to Irish English-speakers. It is possible that Irish-English speakers have less assimilation of nasality than American speakers. Trindade et al. (1997) hypothesised that this could account for lower nasalance scores in Brazilian Portuguese speakers when reading nasal passages. However, if comparisons between Brazilian Portuguese speakers and other studies (Table 3.5) are made, it is evident that the Brazilian Portuguese speakers had lower mean nasalance scores than other speakers on non-nasal passages. Trindade et al. (1997) stated that it is unlikely that assimilation of nasality is a factor in Brazilian Portuguese. However, as the nasalance scores for Irish English-speakers and American speakers differ on speech samples containing nasal consonants but do not differ on non-nasal speech samples, assimilation is a likely explanation.

Nasalance scores on Low Pressure consonant sentences are compared to two other studies which included a Low Pressure consonant speech sample (Haapanen, 1991a; Trindade et al., 1997) (Table 3.6).

Table 3.6 Normal Nasalance scores for sentences containing low pressure consonants and devoid of nasal consonants and high pressure consonants.

Participants	**Investigator**	**N**	**Speech Sample**	**Mean NM**	**SD**
Finnish Adults/ Children	Haapanen, 1991a	38	1 Sentence	12%	7%
Brazilian Portuguese Children	Trindade et al., 1997	20	Brazilian Portuguese Zoo(2)	10%	5%
Irish Children	*Sweeney et al.,*	*70*	*2 Sentences*	*16%*	*6%*

The mean nasalance score for this speech sample ranged from 10% (Trindade et al., 1997) to 16% (present study). Furthermore, there was variation in differences between mean nasalance scores for High Pressure consonant sentences and mean nasalance scores for Low Pressure consonant sentences in the three studies. In the Finnish study the mean nasalance score is 2% higher (14%) for the High Pressure consonant sentences than for the Low Pressure consonant sentences (12%), whereas in the Irish-English study the mean nasalance score is 2% lower for the High Pressure consonant sentences than for the Low Pressure consonant sentences. In the Brazilian Portuguese study there was only a 1% difference with a higher nasalance score on High Pressure consonant sentences (10%). It is difficult to know from the few studies in the area if this is an important difference clinically. It is possible that differences can be attributed to the different languages and sample size used in the studies. Interestingly, standard deviation scores were similar in all three studies. Further studies of the difference between High Pressure consonant sentence scores and Low Pressure consonant scores for English would help ascertain if the difference in scores is due to variations in language or dialect.

The mixed consonant speech sample contained 16% nasal consonants as compared to 11% on the total speech sample. As expected, the nasalance score for the mixed consonant sentence category is higher due to the increased number of nasal consonants.

The mean nasalance score for the Nasal Sentence in the present study was lower than the mean nasalance scores for a Nasal Passage in previous studies (Fletcher et al., 1989; Seaver et al., 1991; Haapanen, 1991a; Anderson, 1996; van Doorn & Purcell 1998) (Table 3.7). This may be due to methodological differences, including the use of different languages, different sample size and different length of utterance in each study.

Table 3.7 Normal nasalance scores for sentences containing nasal consonants.

Participants	**Investigator**	**N**	**Speech Sample**	**Mean nm**	**SD**
American Children	Fletcher et al., 1989	117	Nasal Passage	61%	7%
American Adults	Seaver et al., 1991	148	Nasal Passage	62%	6%
Finnish Adults/ Children	Haapanen, 1991a	12	1 Nasal Sentence	69%	8%
Puerto Rican Spanish Female/Adults	Anderson, 1996	40	Spanish Nasal sentences	62%	8%
Brazilian Portuguese Children	Trindade et al., 1997	20	Brazilian Port. Nasal	51%	6%
Australian Children	van Doorn & Purcell, 1998	245	Nasal Passage	60%	8%
Irish Children	*Sweeney et al.*	*70*	*1 Nasal sentence*	*51%*	*7%*

The Nasal Passage from the Nasometry package was the speech sample analysed in the American (Fletcher et al., 1989; Seaver et al., 1991) and Australian (van Doorn & Purcell, 1998) studies. This passage contains 35% nasal consonants. Despite a higher percentage of nasal consonants in the Irish study (50% nasals), the nasalance scores were lower than the scores from the American and the Australian studies. To some extent this supports the notion of increased perceived nasality and increased assimilation of nasality in American speakers compared to Irish English-speakers (Wise, 1957). Further studies assessing perception of nasality and assimilation of nasality in normal speakers with American, Australian and Irish English would be helpful in comparing the nasalance scores. Interestingly the Brazilian Portuguese passage had a higher proportion of nasal to oral consonants (66%) compared to the Nasal sentence in the present study (50%), but similar nasalance scores were found in the two studies (Table 3.7). The influence of language variation must be considered here. The highest nasalance score on a Nasal sentence was found in the Finnish study which contained 57% nasal consonants. The small number of normal speakers tested in that study makes it difficult to compare scores reliably. Anderson (1996) used 5 nasal sentences, each containing a different proportion of nasal to oral consonants. She reported that, in general, sentences with a higher proportion of nasal consonants had the highest group means, but that there was no one-to-one correspondence between the proportion of nasal consonants and mean nasalance scores. This lack of one-to-one correspondence is supported by comparison of all the studies reviewed in this paper.

3. 5 Limitations of the Normative Study

A limitation of the present study was the unequal number of sentences used in each speech category. This was unavoidable since the aim of the study was to obtain normal nasalance scores for a speech sample that contained all consonant types and therefore was representative of normal speech. The speech sample was designed to be part of a perceptual profile for assessment of nasality and nasal airflow in order to assess phonemes that are vulnerable to velopharyngeal dysfunction (Phillips, 1986; Sell et al., 1994). It was also important to ensure that this same speech sample could be used in a further study comparing perceptual and instrumental assessment. Furthermore, this same speech sample could be used for systematic analysis of consonant errors since each sentence assessed an individual vulnerable consonant.

In order to ensure that the Total Test sentences had 11% nasal consonants (representative of normal conversational speech), only one Nasal sentence was included in the speech sample. This is acknowledged as a limitation of this normative study. Nichols (1999) found that there was a reduction in the stability of measurements when the number of sentences used in the sample was reduced. However, he pointed out that in a study with many participants, the overall means may be reliable, but assessments of individuals may not. Hence, the speech sample used in this study may have been reliable for research but may not be reliable for clinical assessments. Further investigations should include a larger Nasal sample.

3. 6 Conclusions

The present study provides normative nasalance data for English-speaking Irish children, which has not previously been established. Normative nasalance scores for sentences containing all consonant types, High Pressure consonant and Low Pressure consonant sentences (both devoid of nasal consonants) have been reported. Although a mean nasalance score for the Nasal sentence has been found, its clinical value will need to be determined. This study indicated that there was no difference in mean nasalance scores between girls and boys. The differing results of this study and previous studies indicate the need to obtain norms for each population, before the Nasometer can be used reliably for clinical purposes. The findings of this study will be used as baseline normal nasalance values in order to compare perceptual and instrumental measurements of children with abnormal speech.

CHAPTER 4

PERCEPTUAL FRAMEWORK

Introduction

The need for definition of the terms used to describe the speech problems associated with cleft palate, non cleft velopharyngeal dysfunction and nasal obstruction has been emphasised in the literature (Bzoch, 1989; Kent, 1996). Working definitions of terms were devised for the present study in order to develop a perceptual assessment of nasality and nasal airflow errors. These definitions are presented in the first section of this chapter. Scaling techniques have been used traditionally in the assessment of hypernasality (Henningsson & Hutters, 1997). However, points on these scales have not been clearly defined (Wirz & Mackenzie, 1995). A novel type of scale has been devised for the present study. This scale describes speech symptoms in terms of the nature of the errors perceived by the listener and is therefore descriptive in nature. Section 4.2 is a description of the Perceptual Profile. The speech sample has been developed for use with it, taking account of the phonetic makeup of the utterances and the importance of assessing different levels of speech (i.e., word level, sentence level, automatic speech and conversational speech) (Sell et al., 1994; D'Antonio & Scherer, 1995; Henningsson & Hutters, 1997). Articulatory errors and voice quality were recorded as possible factors which influence assessment of nasality and nasal airflow errors. The importance of evaluating reliability of assessment procedures has also been highlighted (Kearns & Simmons, 1988; Cordes, 1994). The final sections of this chapter outline the different studies of reliability of the Perceptual Profile, which have been undertaken. These studies included intra-rater reliability within listeners, evaluating listening conditions and use of anchor stimuli, and inter-rater reliability between listeners. Reliability studies included the calculation of percent agreement, as well as kappa analyses, which takes into account chance agreements (Kreiman et al., 1993).

4. 1 Definition of Terms

Perceptually based identification of disorders of nasality present many problems for the team working with speech problems associated with cleft palate and non cleft velopharyngeal dysfunction. One of the basic problems, according to Bzoch (1989), is that the concept of nasality must be restricted by definition. He stated that nasality can only be reliably identified or logically discussed in clinical or speech science research, if the term is defined and strictly limited in use. The following proposed framework attempts to address this requirement.

Nasality and nasal airflow errors are general terms and are defined as follows:

Nasality refers to perceived nasal resonance during production of multi-segmental units resulting from the coupling of the oral and nasal resonating cavities. In English, nasality, or nasalization as it is sometimes called (Kittelson et al.,1983; Grunwell & Harding, 1996), typically occurs on nasal consonants /n, m, ng/ and on adjacent vowels. Laver (1980) states that the vital factor in inducing nasal resonance is the ratio of nasopharyngeal port size to the oropharyngeal port size. The typical acoustic characteristic associated with nasality is a reduction of intensity of the first formant (Swartz 1971; Moon, 1993). Other acoustic features which have been associated with nasality include the presence of antiresonances, the presence of extra resonances within the spectrum and a shift in the centre frequency of the formants (Moon, 1993). Nasality in the present study refers to normal nasal resonance, and also includes abnormal resonance of hypernasality, hyponasality and cul de sac resonance.

Nasal airflow errors refer to the inappropriate escape of air through the nasal cavity during production of voiced and voiceless pressure consonants. Kittelson et al. (1983) distinguish between the nasal frictional quality of nasal airflow errors as opposed to the resonant quality described in nasality. Nasal airflow errors are segmental features associated with atypical oro-nasal structure and/or function (Grunwell & Harding, 1996). Nasal airflow errors can be inaudible (Sell et al., 1994; Sell et al., 1999), in which instance they do not interfere with the speech signal (Bzoch, 1989). Therefore, inaudible nasal airflow errors are not included in this definition of nasal airflow anomalies as these definitions form the basis of the Perceptual Profile. Trost (1981) reported that audible nasal airflow errors have been described perceptually in many different ways, including nasal snort, hissing, nasal emission and nasal turbulence. Nasal airflow errors may result from incomplete closure of the velopharyngeal sphincter during sound production or an anterior oronasal fistula. Kittleson et al (1993) report that these sounds are characterised by the release of noise energy through the nasal cavity. The physical correlates of nasal airflow errors have not been well described; however, an increase in the rate of nasal flow through the nose during sound production would be expected (Warren, 1995, personal communication). Four different categories of nasal airflow errors will be presented below in section 4. 2. 1.

Nasality and nasal airflow errors can co-exist; a speaker may present with both hypernasality and nasal airflow errors. Alternatively, nasality and nasal airflow errors can exist independently. For example, a speaker can have abnormal nasality with no auditorily detectable nasal airflow errors, or normal nasality with nasal airflow errors.

4. 1. 1 Nasality

There are three main types of nasality requiring suprasegmental analysis:

> *Hypernasality* is the occurrence of excessive nasal resonance perceived during speech production. It results from a coupling of oral and nasal resonating cavities when the velopharyngeal sphincter is in an open position. According to Laver's (1980) classification, hypernasality occurs when there is an increase in the ratio of nasopharyngeal port to oropharyngeal port size. Hypernasality is perceived during multisegmental units that normally do not have perceived nasality or its acoustic correlates. The acoustic correlates of hypernasality include a reduction of intensity of the first formant, and sometimes the presence of antiresonances and/or extra resonances within the spectrum and a shift in the centre frequency of the formants on the vowels and voiced sonorants throughout the utterance. The

degree and consistency of hypernasality can vary.

Hyponasality is the reduction or absence of expected nasal resonance associated with nasal consonants and vowels adjacent to nasal consonants in English. This is usually due to the reduction of the size of the velopharyngeal port and/or nasal airway, resulting in a reduction or absence of the resonance within the nasal cavity. According to Laver's (1980) classification, there is a decrease in the ratio of the nasopharyngeal port size to the oropharyngeal port size. He states that the acoustic correlate of hyponasality is the minimisation of the acoustic characteristics of nasality (i.e. the intensity of the first formant is not reduced, the presence of antiresonances and/or extra resonances within the spectrum and a shift in the centre frequency of the formants are not evident for nasal consonants and adjacent vowels). The degree and consistency of hyponasality can vary.

McWilliams et al. (1990) define *Cul de Sac Resonance* as the coupling of a closed nasal resonating cavity to the oral resonating cavity. It is thought to be due to blockage of the anterior section of the nasal cavity. Auditorily, there is a subtle difference between hyponasality and *Cul de Sac* resonance. Although this difference is not defined perceptually or acoustically in detail, *Cul de Sac* resonance has been described as a variation of hyponasality which differs only in the place of nasal obstruction (McWilliams et al., 1990).

4. 1. 2 Nasal Airflow

According to the above definition, nasal airflow does not exist in normal English speech, but is associated with deviant speech production. A segmental analysis is required for assessment of nasal airflow problems. Working definitions of four different nasal airflow errors are presented here and will be discussed in relation to other definitions presented in the literature.

Nasal emission is audible escape of air from the nasal cavity accompanying production of oral pressure consonants. It is due to incomplete closure of the velopharyngeal sphincter and/or the presence of a fistula in the palate. A frictional sound is produced when the airflow is sufficiently strong and the constriction is sufficiently narrow to create noisy random vibrations in the airstream (Borden & Harris, 1980). In the case of nasal emission, the frictional sound is a result of strong airflow passing through the nasal cavity which has sufficient constriction within it. Nasal emission has a frictional but no turbulent or snorting quality. During consonant production there is oral and nasal release of air, as the nasal emission accompanies the consonant (Sell et al., 1994). In the extended International Phonetic Alphabet (IPA) (Duckworth et al., 1990), the diacritic representing nasal emission is [͊] e.g. [s͊].

A *nasal fricative* also has the frictional sound produced by air passing through the nasal cavity when the velopharyngeal sphincter is open. However, there is complete or almost complete stricture in the oral cavity resulting in no audible oral release. Hence, the nasal fricative replaces a consonant, unlike nasal emission which accompanies the consonant. It has been described as the realisation of a target oral consonant as a nasal, with airflow through the nasal cavity creating friction (Eurocleft Speech Group, 1993; Grunwell & Harding, 1996). Nasal fricatives can be transcribed as [m̥ᶠ] or [n̥ᶠ] depending on location of oral stricture (Sell et al., 1994).

Nasal turbulence is defined as a 'snorting' or turbulent noise resulting from the approximation but inadequate closure of the superior border of the velum and the posterior pharyngeal wall (Duckworth et al., 1990). This has been described by Kummer et al. (1992) as a 'nasal rustle' resulting from friction produced when an airstream passes through a small velopharyngeal gap. They state that this sound is louder and more distorted than nasal emission. Trost (1981) reported that the velar activity can be seen radiographically as a velar flutter. Hence, the nasal turbulence may be a type of velar trill, where there is rapid light contact between the velum and pharyngeal walls (Riski, 1999, personal communication). Like nasal emission, nasal turbulence accompanies other consonants resulting in oral and nasal release of air. The extended IPA diacritic for nasal turbulence is [͌] e.g. [s͌] (Duckworth et al., 1990).

A *velopharyngeal fricative* is a 'snorting' or turbulent noise, which is used as a substitution for another consonant. It has been described by Duckworth et al. (1990) as the approximation but inadequate closure of the superior border of the velum and the posterior pharyngeal wall. There is complete or almost complete oral stricture during sound production with no audible oral release. As in nasal turbulence, the resulting sound may be due to strong friction as air passes through the narrow velopharyngeal port, or the velar trill. Whereas nasal turbulence accompanies consonants (Sell et al., 1994), the velopharyngeal fricative acts as a sound substitute (Duckworth et al., 1990). The extended IPA transcription for a velopharyngeal fricative is [ʩ] (Duckworth et al., 1990).

4. 1. 3 Discussion of Nasality and Nasal Airflow Errors Definitions

The literature indicates the complexities and controversies regarding the definitions of nasality and nasal airflow errors in speech. One of the main difficulties is the variation of terms used to describe the speech characteristics (Laver, 1980; Bzoch, 1989). As Laver (1980) points out, the terms used to describe nasality have been used interchangeably, such as nasal voice, denasal voice, hypernasality, hyponasality and rhinolalia. However, it is now recognised that these terms are subcategories of the generic term 'nasality'. There are also some controversies regarding definitions of nasal airflow errors in the literature. Grunwell and Harding (1996) define audible nasal escape as the presence of auditorily detectable airflow through the nose accompanying the production of target oral consonants, specifically obstruents. Nasal emission frequently accompanies voiceless pressure consonants. Sell et al. (1994) report that nasal emission can accompany or replace a consonant. Other reports state that when nasal emission replaces a consonant, nasal fricatives are produced (Harding & Grunwell, 1998). In the classification in the present study, nasal emission is defined as airflow accompanying a consonant, whereas the term nasal fricative is used when nasal emission replaces a consonant. Hence, nasal emission and nasal fricatives have the same velopharyngeal position during consonant production, but different oral articulatory patterns. For example, an /s̃/ is produced with lateral alveolar-lingual contact with a central oral airstream and accompanying nasal emission, whereas a /n̥/ sound has complete alveolar-lingual closure and the airflow is directed through the nasal cavity.

Nasal fricatives have been classified as active or passive (Harding & Grumwell, 1998; Sell et al., 1999). According to the Harding and Grunwell (1998) definition an active nasal fricative is produced by actively directing the airflow nasally and stopping oral airflow with the lips or the tongue. "This strategy is developed in order to signal the fricative nature of the intended target fricative consonants" (Harding and Grunwell; p.334, 1998). They define passive nasal fricatives as an unreleased /s/ which is doubly articulated with a lowered voiceless nasal [(s)͡n̥]. For passive nasal fricatives the oral airflow is not stopped, and the air passively escapes into the nasal cavity. The distinction between active and passive nasal fricatives may not be perceptually distinguishable but may be identified by pinching the nose and occluding nasal airflow. If the nasal fricative is passive then the target phoneme will be produced. Because it is difficult to distinguish active and passive nasal fricatives perceptually, this distinction was not included in the present Perceptual Profile.

Various terms have been used to describe nasal turbulence. Duckworth et al. (1990) refer to it as a snorting sound, while Kummer et al. (1992) label nasal turbulence as a 'nasal rustle'. The term nasal turbulence is perhaps a misnomer for two reasons. Firstly, the noise is produced at the velopharyngeal port and not in the nasal cavity. Secondly, it is ill-understood how this sound is produced. It may be a result of sufficient narrowing in the velopharyngeal port to produce the snorting sound (Duckworth et al., 1990; Kummer et al., 1992), or it may be a result of a velar trill (Riski, 1999, personal communication). If the latter explanation is accepted then the term velopharyngeal trill may be more accurate. Nasal turbulence has been defined as a more severe form of nasal emission (McWilliams et al., 1990; Sell et al., 1994; Grunwell & Harding, 1996). In contrast, Kummer et al. (1992) reported that nasal emission is usually associated with a wide velopharyngeal gap and that nasal turbulence is associated with a small velopharyngeal gap. As the precise articulatory production of this sound is not yet confirmed and as the term nasal turbulence is widely accepted as a recognised characteristic of cleft palate speech (Sell et al., 1994; 1999; Wyatt et al., 1996; CSAG, 1998), it seems reasonable to use nasal turbulence in the auditory perceptual framework for this study. Trost (1981) also described an audible resistance to nasal airflow that is intra-nasal. This resistance causes a type of nasal turbulence that will be referred to as *intra-nasal turbulence* in this study. This can be associated with nasal emission and nasal fricatives, when the turbulent sound is produced within the nasal cavity, at the site of a nasal obstruction, and not at the velopharyngeal sphincter. This type of turbulence should be distinguished from nasal turbulence.

Another difficulty is that there is controversy regarding how specific speech characteristics are produced. Various theories have been postulated regarding the production of hypernasality. Laver (1980) states that the ratio of the velopharyngeal port is increased in relation to the oropharyngeal port. Warren (1993b) states that the timing of velopharyngeal closure may be important in producing hypernasality. With regard to nasal turbulence, as indicated above, the sound may be a result of sufficient narrowing in the velopharyngeal port to produce the snorting sound (Duckworth et al., 1990; Kummer et al., 1992), or it may be a result of a velar trill (Riski, 1999, personal communication).

Apart from the problems outlined above, the acoustic correlates of the nasality and nasal airflow errors have not been fully determined. Although it is accepted that the acoustic correlate of hypernasality includes a reduction of intensity of the first formant, the other acoustic features are not consistently present in hypernasal voice. Laver (1980) states that, antiresonances and/or extra resonances within the spectrum are found only on an inconsistent basis and there can be a shift in the centre frequency of the formants on the vowels and voiced sonorants. He stresses that the exact detail of the changes in the acoustic spectrum depends on the configuration of the vocal tract at a precise moment. No acoustic analyses of nasal emission and nasal turbulence have been reported in the literature. Preliminary spectral analyses of participants with nasal emission and nasal turbulence were carried out in the present study. Spectrograms indicated acoustic differences between speakers with nasal emission/nasal fricatives where there was no 'snorting' sound, and nasal turbulence/velopharyngeal fricatives where there was a 'snorting' sound. One participant with nasal emission had high frequency energy between 3000 and 5000hz during the production of the plosive /t/ in the sentence 'Tim had a tart for tea' (Figure 4.1).

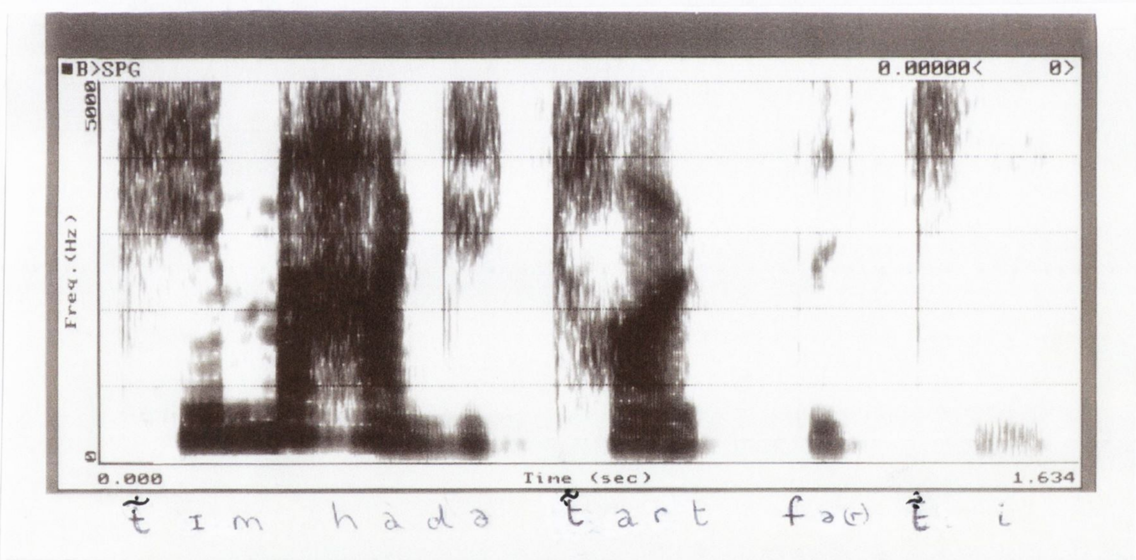


Figure 4.1 Spectrogram of a speaker with nasal emission during the production of the sentence 'Tim had a tart for tea'. Note the high frequency energy between 3000 and 5000Hz associated with the production of /t̃/ with accompanying nasal emission.

Another participant who had nasal turbulence had a band of low frequency energy at approximatly 100 hz during the production of the plosive /t/ in the same sentence (Figure

4.2).

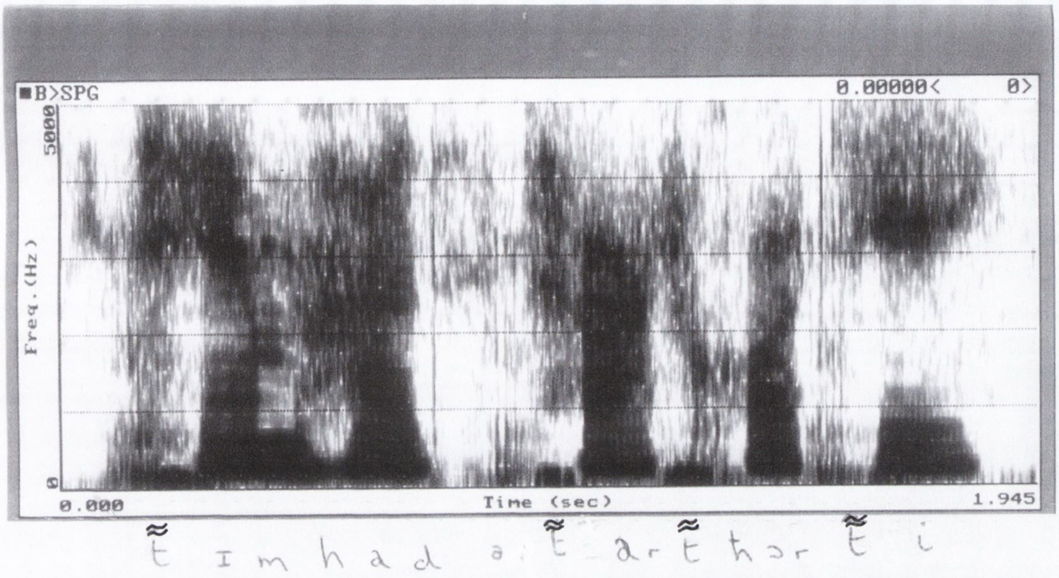


Figure 4.2 Spectrogram of a speaker with nasal turbulence during the production of the sentence 'Tim had a tart for tea'. Note the dark band of low frequency energy (around 100 Hz) associated with the production of /t̃/ with accompanying nasal turbulence.

The speech of eight participants was analysed acoustically. Two participants from each of the four categories of nasal airflow errors were included. The analyses indicated that participants with nasal emission and nasal fricatives had similar acoustic patterns, with varying degrees of high frequency energy between 3000 and 5000 hz. However, participants with nasal turbulence and velopharynegal fricative had varying degrees of low frequency energy at around 100 hz. This low frequency energy may be associated with the snorting sound which is present in these two categories of nasal airflow errors. Although the acoustic analyses in the present study were limited, there are indications that there may be specific acoustic characteristics associated with the two broad categories of nasal airflow errors. These data highlight the need for further detailed acoustic investigations into nasal airflow errors in speech.

A velopharyngeal fricative may act as a substitute for a phoneme (Trost, 1981), as in the case of a phoneme specific velopharyngeal fricative. This has been referred to as "phoneme specific nasality" (Albery, 1989). The use of this latter term, however, may lead to confusion, as it implies a resonance problem as opposed to an airflow problem. Usually a velopharyngeal fricative is used as a substitute for the /s/ and /z/ phonemes. It is usually indicative of velopharyngeal mislearning in the absence of a structural problem. Although Grunwell and Harding (1996) report that this type of airflow problem is rare, its description by other authors (Peterson, 1975; Trost, 1981; Albery, 1989) suggests that it is important to include the velopharyngeal fricative in the development of an assessment profile as it has significant implications for management.[1]

The terms *velopharyngeal insufficiency, velopharyngeal incompetency, velopharyngeal inadequacy* and *velopharyngeal dysfunction* have been used to describe various abnormalities of velopharyngeal function. Loney and Bloem (1987) reported that authors used these terms interchangeably or used one term to describe all types of velopharyngeal malfunction. D'Antonio and Scherer (1995) highlighted the problems of inconsistent use of terms. They stated that imprecise use of terminology can lead to erroneous assumptions regarding a patient, and inappropriate clinical actions. They proposed the following definitions of terms which are now largely accepted in the literature:

Velopharyngeal dysfunction is a generic term which denotes any type of abnormal velopharyngeal function resulting from structural deficits, neurological disorders, faulty learning or a combination of aetiologies. The following terms are included under this broad category: *velopharyngeal insufficiency* occurs where there is a lack of sufficient tissue to effect velopharyngeal closure; *velopharyngeal incompetency* indicates a lack of neuromuscular competency in opening and closing the velopharyngeal sphincter; *velopharyngeal mislearning* refers to maladaptive articulation which is learnt and is not due to structural or neurological aetiologies.

[1] If a patient is using a velopharyngeal fricative as a phoneme substitution, without any other abnormal nasality or nasal airflow, it is suggested that the problem is one of mislearning rather than structural. In such cases a phonological approach to treatment is recommended and the prognosis with speech and language therapy is good (Bzoch, 1989)

4. 1. 4 Summary

The above "working definitions" have been proposed in order to develop a reliable and valid perceptual assessment tool for clinical use. Nasality is defined at a suprasegmental level and includes hypernasality and hyponasality. Nasal airflow definitions are at a segmental level. They are based on descriptions by Grunwell and Harding (1996) and are further developed for use in clinical assessment.

4. 2 Description of the Perceptual Profile

The present definitions form the basis of a new perceptual assessment framework for comparison with data from objective measurements of nasal resonance and airflow. It is hoped that the study of the data from objective measurements and perceptual data may resolve some of the controversies, thereby facilitating a definitive framework of perceptual speech characteristics related to nasality and nasal airflow errors.

The Perceptual Profile is divided into two sections: Nasality, which includes hypernasality, hyponasality and cul de sac resonance and Nasal Airflow errors, which include nasal emission, nasal fricatives, nasal turbulence and velopharyngeal fricatives. Descriptors are used to rate nasality, and each descriptor is identified by a letter. In this way the focus of the scale is on the description of the speech problem.

4. 2. 1 Nasality

Hypernasality (the occurrence of excessive nasal resonance) is rated as present or absent. If present, it is rated in terms of severity as follows:

a) mild, evident but acceptable;

b) mild/moderate, unacceptable distortion evident on high vowels;

c) moderate, evident on high and low vowels;

d) moderate/severe, evident on all vowels and some consonants;

e) severe, evident on all vowels and most voiced consonants.

Hypernasality is rated as consistent if the severity of hypernasality is the same for all levels of speech (i.e., words, sentences, automatic speech and conversational speech) (see speech samples below). Hypernasality is inconsistent when the severity of hypernasality differs across the different speech levels.

Hyponasality (the reduction or absence of expected nasal resonance) is rated as present or absent. If present, it is rated in terms of severity as follows:

a) evident, but acceptable;

b) moderate, reduced nasality on all vowels and some consonants;

c) severe, total denasal production of nasal consonants.

Hyponasality is rated as consistent if the severity of hyponasality is the same for all levels of speech. It is rated as inconsistent when the severity differs across the different speech levels.

Cul de sac resonance (the coupling of a closed nasal resonating cavity to the oral resonating cavity) is rated as present or absent.
(Appendix 13b).

4. 2. 2 Nasal Airflow Errors

Nasal airflow errors are divided into four categories which can exist independently or together. The four categories include:

- nasal emission (audible escape of air from the nasal cavity accompanying production of oral pressure consonants);
- nasal fricatives (the realisation of a target oral consonant as a nasal with airflow through the nasal cavity creating friction);
- nasal turbulence (a velopharyngeal frictional noise resulting from approximation but inadequate closure of the velopharyngeal sphincter); and
- velopharyngeal fricatives ('snort' which is a substitution for another sound, resulting from approximation but inadequate closure of the velopharyngeal sphincter and complete oral stricture).

If a nasal airflow error is identified, it is classified according to the above definitions and rated in the following manner:

weak/strong - weak nasal airflow refers to low intensity of audible nasal airflow on segments. Strong nasal airflow refers to high intensity of audible nasal airflow;

frequent/infrequent - frequency of nasal airflow is calculated during sentence repetition and automatic speech, according to the percentage of phonemes exhibiting nasal airflow errors in the speech sample. Ninety-two phonemes have been identified as possible targets for nasal airflow. Frequent nasal airflow indicates more than 10% of target phonemes with nasal airflow. Infrequent nasal airflow indicates up to 10% of target phonemes with nasal airflow;

consistent/inconsistent - the error is rated as consistent if the strength and/or frequency of occurrence is the same across all the levels of speech (words, sentences, automatic and conversational speech). If there is a difference in the strength and/or frequency of nasal airflow errors across the different speech levels, then it is rated as inconsistent;

phoneme specific indicates that the nasal airflow error is evident on specific sounds only.
(Appendix 13b)

4. 2. 3 Speech Stimulus

Speech stimuli are presented in sections representing levels of speech: words, sentences, automatic speech and conversational speech (Appendix 10)[2]. The test words were adapted from the stimulus, which was to be used later in an instrumental assessment.[3] The test sentences were adapted from the speech sample used during the development of the GOS.SP.ASS (Sell et al., 1994). Automatic speech included a nursery rhyme "Jack and Jill", and counting from 1 to 20 and 60 to 70. Conversational speech included a minimum of two minutes spontaneous speech in response to simple questions such as

---

[2] All high pressure consonants which were vulnerable to nasal airflow errors were underlined to aid in counting the frequency of nasal airflow errors.

[3] This speech sample was to be used for aerodynamic assessment using the NORS. However, due to problems in reliability and validity, the NORS was not used in the instrumental study. The Test Words were subsequently changed for the instrumental assessment so that the same word stimuli could be used for perceptual and aerodynamic assessments.

"tell me your name", " where do you live?", "how old are you?" and "tell me about your brothers and sisters".

During production of the speech stimuli the listener rates nasality and nasal airflow errors under each speech stimulus. Errors of articulation are also noted in each section. The voice quality is recorded as normal, dysphonic, or with reduced volume.

The following sections describe the reliability studies that were carried out using the Perceptual Profile and speech stimulus.

4. 3 Reliability of the Perceptual Profile

Four reliability studies were carried out using the Perceptual Profile over a two year perios. Three different but related intra-rater reliability studies were completed over a ten month period, by the same rater. In each of the three intra-rater studies, another therapist not involved in the study rearranged the order of data presentation for the repeat analyses. Following this a fourth study evaluated inter-rater reliability. The first study (study 1: intra-rater reliability using live assessments versus audio recorded assessments) compared ratings of nasality and nasal airflow errors during a live assessment session with ratings from an audio recording of the same session. All participants were assessed over an eight week period. Three weeks following the completion of live data collection, the audio recorded speech samples were analysed again.

The second study (study 2: intra-rater reliability of the Perceptual Profile using audio recorded speech samples) compared ratings from the audio recordings used in study 1 with ratings of the same recording analysed ten weeks later.

The third study (study 3: intra-rater reliability using *anchor stimuli*) was carried out five months after completion of study 2. The audio recorded speech samples used in study 2 were analysed again using anchor stimuli. After a delay of ten weeks later, the audio recorded data was analysed again using anchor stimuli. The ratings of the two analyses were compared.

The fourth inter-rater study compared ratings of three different raters using the Perceptual Profile. Audio taped speech samples from the intra-rater study and the pilot study were included (Appendix 12). Anchor stimuli were also used. This study was carried out three months after the intra-rater reliability studies.

4. 4. Study 1: Intra-Rater Reliability Using Live Versus Audio Recorded Assessments

4. 4. 1 Methodology

Participants

12 children between ages 3;4 years and 13;10 years attending the Cleft Palate Unit were assessed. Children with cleft palate, velopharyngeal dysfunction, or nasal obstruction, with or without syndromes, were included. All participants presented with abnormal nasality and/or nasal airflow problems, with or without articulatory errors.

Children were excluded if there was evidence of any of the following: severe dyspraxia/dysarthria; learning disability (greater than mild); bilateral hearing loss above 45 db.; nasal blockage associated with an upper respiratory infection; moderate to severe hoarseness of voice; mixed nasal resonance; and an inability to complete the assessment protocol due to behavioural problems or poor co-operation.

Speech sample

The child produced the speech stimuli at each speech level (words, sentences, automatic and conversational speech).

1. The child repeated the test words after the rater

2. The child repeated 16 test sentences after the rater
3. Automatic speech was assessed by asking the child to recite a nursery rhyme, count from 1 to 20 and from 60 to 70.
4. Each child was engaged in two minutes of conversational speech.

(Appendix 10)

Procedures

Speech was rated during a live assessment session in a quiet room, with the rater sitting opposite the child. The rater listened to the child's speech and rated nasality and nasal airflow errors according to the Perceptual Profile at word, sentence, automatic and conversational levels of speech. During the assessment of nasality and nasal airflow, articulatory errors were noted on the assessment sheet (Appendix 13b).

During the live assessment, the speech samples were audio recorded by the rater using a Sony Scoopman digital audio tape recorder and a Sony condenser microphone. Positioning of the microphone was previously tested by the examiner and a sound engineer. Both parties listened to recorded speech samples with the microphone in different positions in relation to the speaker, and at varying decibel recording levels. It was noted that optimal audio recording was made when the microphone was placed at the level of the child's mouth, slightly to one side and approximately four inches away. A recording level between 25db and 28db was noted to be adequate.

Three weeks following the live analyses, the the audio recordings were randomly arranged by another therapist who was not involved in the study. Audio playback was carried out in a quiet room at a decibel level of 20db to 25db, with the use of earphones. The rater listened to each speech level and rated each speech level accordingly on the Perceptual Profile Sheet (Appendix 13b).

Reliability was statistically calculated using kappa analysis and the percent of agreement between ratings. The kappa statistic relates the actual measure of agreement obtained with the degree of agreement that would have been obtained had the diagnosis been made at random (Bulman & Osborn, 1989). A weighted kappa was used in this study as this takes into account 'near misses', where close agreement on ratings is credited (Bulman & Osborn, 1989). Expected agreement, i.e. agreement that would be expected by chance, is also presented. The kappa statistics are presented where 1 represents perfect agreement, over 0.75 indicates excellent agreement, between 0.4 and 0.75 represent fair to good agreement beyond chance, and below 0.4 represents poor agreement (Fleiss, 1981). Based on previous studies, 90% to 100% agreement was considered excellent, 80% to 89% was good and 75% to 80% was acceptable (Lohmander-Agerskov, 1996; Watterson et al., 1998a). For the purpose of the present study, the term 'reliability' will be used in a broad sense to include percent agreement and kappa scores.

4. 4. 2 Results of Reliability between Live and Audio Recorded Assessments

Reliability coefficients and percent agreements between live and audio analysis were calculated for the following speech parameters: nasality; consistency of nasality; consistency, strength, frequency and phoneme specificity for each of the four nasal airflow errors (Table 4.1). Percentage of agreement between live and recorded assessments ranged from 75 % to 96% indicating good to excellent agreement between assessments. Excellent reliability using kappa scores were found on ratings of nasality, consistency of nasality, and all aspects of nasal emission. Kappa scores for velopharyngeal fricative consistency ranged from -0.1 to 0.8. Low kappa scores were found on nasal fricative and velopharyngeal fricatives.

Table 4.1 Intra-rater reliability (percent agreement, expected agreement and kappa scores) for ratings of nasality and nasal airflow errors during live and audio recorded analyses.

Speech Parameter	% Agreement	% Expected Agreement	Kappa
Nasality	89	58	0.7
Consistency	87	61	0.7
Nasal Emission			
strength	91	76	0.6
consistency	96	76	0.8
frequency	92	79	0.6
phoneme specific	92	67	0.7
Nasal Fricative			
strength	88	82	0.3
consistency	87	82	0.3
frequency	87	82	0.3
phoneme specific	83	72	0.4
Nasal Turbulence			
strength	83	72	0.4
consistency	83	63	0.5
frequency	79	61	0.5
phoneme specific	83	72	0.4
Velopharyngeal Fricative			
strength	83	72	0.4
consistency	77	79	- 0.1
frequency	75	72	0.1
phoneme specific	83	79	0.2

4. 4. 3 Live versus Audio Recorded Assessments - Discussion

Results of the intra-rater reliability study are discussed for each speech parameter.

Nasality

Good reliability was found between intra-rater ratings of nasality during the live assessment and the audio recorded assessment of the live session (percent agreement was 89% and the kappa was 0.7). In all but one case, agreement differed by only one point (Table 4.2).

Table 4.2 Agreement of intra-rater ratings of nasality during the live and audio recorded analyses.

	Nasality - Audio				
Nasality-live	**0**	**1**	**2**	**3**	**4**
0	3	0	1	0	0
1	1	1	0	0	0
3	0	0	1	2	1
4	0	0	0	0	1

Nasal emission

Results indicated good to excellent reliability for ratings of nasal emission - strength, frequency, consistency and phoneme specificity (percent agreement ranged from 91% to 96% and the kappa scores ranged from 0.6 to 0.8).

Nasal fricative

Results for ratings of nasal fricatives varied, with good agreement (83%) and poor kappa scores (0.3 to 0.4). The small number of children who presented with this consonant error may explain the high percent agreement and low kappa scores; therefore, the percentage score may be a more accurate indication of agreement than the kappa score. Three participants were reported to have a nasal fricative. There was complete agreement on the ratings of one participant. The second participant was reported to have a velopharyngeal fricative in both live and audio analysis. However, during the live assessment it was reported that he occasionally used a nasal fricative, as well as frequent use of a velopharyngeal fricative. The nasal fricative was weak and infrequent and was not perceived on audio tape. The third participant was reported to have a velopharyngeal fricative during the live assessment, and a nasal fricative during the audio assessment. As the velopharyngeal fricative was weak, it may have been detected as a nasal fricative on audio tape. In this study there was confusion between the perception of weak nasal fricatives and weak velopharyngeal fricatives. This may have been a perceptual error, or it may have been due to the different listening conditions.

Nasal turbulence

Kappa scores for ratings of nasal turbulence were moderate (0.4 to 0.5), while the percent agreement was good (79% to 83%). Detailed analysis of the data indicated that there were two cases where there were discrepancies between the presence and absence of nasal turbulence. One participant was reported as having nasal emission and no nasal turbulence during the live assessment. However, on audiotape, weak infrequent inconsistent nasal turbulence was detected. Because this turbulence was mild, it may not have been detected during the live session. However, the second participant was reported as having weak nasal turbulence and a weak velopharyngeal fricative during the live assessment. On audiotape analysis she was reported to have a nasal fricative with no nasal turbulence or velopharyngeal fricative. The discrepancy in ratings of the two participants may be due to the mild borderline nature of the problem. McWilliams and Phillips (1979) reported that mild disorders of resonance are often the most difficult to evaluate and that tape recordings can 'lighten' mild speech disorders.

Velopharyngeal fricative

Kappa scores for velopharyngeal fricatives were the lowest of all speech parameters. This may be explained by the small number of occurrences of velopharyngeal fricatives and the confusion with nasal fricative as described above.

4. 4. 4 Summary

Overall results indicate good reliability between the two testing situations. Some discrepancy in results was explained by the different listening conditions. A number of situations occurred where there was confusion between a nasal fricative and a velopharyngeal fricative. This suggests the need to consider reducing the scale to include one category for nasal fricative and velopharyngeal fricative. This possibility will be reviewed further, following comparisons of perceptual ratings of nasal airflow errors and instrumental assessments of airflow. It is evident from the results that weak nasal airflow problems lead to confusion in categorisation of airflow problems. This will be discussed further in the Section 4.10 in the light of the inter-rater reliability results.

4. 5 Study 2: Reliability Of Intra-Rater Ratings Using Audio Taped Speech Samples

4. 5. 1 Methodology

The participants and speech samples used in the first study were included in study 2. In preparation for the repeat analysis, another therapist not associated with the study, randomly arranged the order of presentation. The listening task was repeated ten weeks after the first audio analyses, which was undertaken for study 1. Results from the first audio taped analyses were then compared to the second audio taped analyses. Statistical analysis of reliability was carried out using percent agreement and kappa correlations.

4. 5. 2 Intra-rater Reliability Results

Overall scores in the intra-rater reliability study using audio tape analysis were good to excellent using the above criteria (Fleiss, 1989). The percentage of agreement ranged from 83% to 100% and kappa scores ranged from 0.4 to 1 (Table 4.3). There was excellent agreement on nasal emission, nasal fricatives, and strength and consistency of velopharyngeal fricatives. Good agreement was found on nasality. Moderate agreement of ratings was found on nasal turbulence - strength and phoneme specificity.

Table 4.3 Intra-rater reliability (percent agreement, expected agreement and kappa scores) for perceptual ratings of nasality and nasal airflow errors using the weighted kappa.

Speech Parameter	% Agreement	% Expected Agreement	Kappa
Nasality	88	65	0.7
Consistency	83	59	0.6
Nasal Emission			
Strength	91	76	0.6
Consistency	91	83	0.5
Frequency	92	67	0.8
Phoneme specific	92	67	0.8
Nasal Fricative			
Strength	92	79	0.6
Consistency	96	76	0.8
Frequency	96	76	0.8
Phoneme specific	92	67	0.8
Nasal Turbulence			
Strength	83	72	0.4
Consistency	83	61	0.6
Frequency	92	62	0.8
Phoneme specific	83	72	0.4
Velopharyngeal Fricative			
Strength	100	72	1
Consistency	92	79	0.6
Frequency	88	75	0.5
Phoneme specific	88	76	0.5

4. 5. 3 Intra-rater Reliability Discussion

Nasality

Results indicated good intra-rater reliability on the nasality scale (percent agreement of 88% and kappa score of 0.7). This compares well to agreement in previous studies (Table 4.4). Consistency of nasality showed good agreement (83% agreement and a kappa score of 0.6).

Table 4.4 Previous studies indicating intra-rater reliability using correlational analyses. The reliability results from the present study are presented.

Study	**Reliability Analysis**	**Scale**	**Result**
Ramig (1982)	Pearson correlation	7 points	0.55 - 0.90
Sell & Grunwell (1990)	Spearman Rank correlation	5 points	0.76
Nellis et al. (1992)	Pearson correlation	6 point	0.75
Hardin et al. (1992)	Pearson correlation	5 points	0.8 - 0.9
CSAG (1998)	Spearman Rank		1
Pinborough-Zimmerman (1999)	Pearson	6 points	0.79 - 0.91
Present Study	Percent agreement Kappa score	5 points	88% 0.7

The correlational results in previous studies only indicate that the rater rated each sample in a similar manner. Furthermore, strong correlations do not always indicate good reliability (Watterson et al., 1998b). Unfortunately, none of the studies reported percent agreement and/or kappa scores for reliability of nasality. The rigorous statistical analysis in the present study indicates good reliability.

There was no difference between reliability scores in the assessment of nasality for live and audio recorded assessments compared to the reliability scores for audio and audio recorded assessments. The present results would support previous findings by Moller and Starr (1984) who reported no significant difference between live analysis and audio analysis. However, it is possible that there was some degree of habituation in the second study. Although there was a time lag of ten weeks between the two audio recorded analyses, the samples had been analysed three times and the rater was more familiar with the data during the second audio recorded analyses.

Nasal emission

The percent agreement and kappa scores for nasal emission ranged from good to excellent in this study. There was 91/92% agreement on ratings of all parameters of nasal emission, while kappa scores ranged from moderate (0.5 for consistency) to excellent (0.8 for frequency and phoneme specificity). The CSAG (1998) study reported intra-rater reliability for nasal emission, using a shortened scale. This study reported variability in intra-rater reliability, with one rater demonstrating excellent intra-rater reliability (kappa score of 1.0), and the second rater demonstrating weak intra-rater reliability (kappa score of 0.42). The present findings compare well with the CSAG results, considering the detail of the perceptual scale used in this study. The results indicate that the profile has good intra-rater reliability for assessment of nasal emission.

A comparison of results from study 2 (repeated audio taped analyses) and study 1 (live versus audio taped analyses) indicated minimal differences in reliability. One would have expected improvements in reliability in study 2, as the same listening conditions were used, and there was some familiarity with the data. For consistency of nasal emission there was greater reliability in study 1. An agreement of 96% and a kappa score of 0.8 were found in study 1, while an agreement of 91% and a kappa score of 0.5 were found in study 2. Otherwise, the results from the two studies were similar. This again supports the results of Moller and Starr (1984) who found no significant differences in assessment results for different listening conditions.

Nasal fricative

Results indicated good to excellent intra-rater reliability on assessment of nasal fricatives. Agreements ranged from 92% to 96%, while kappa scores ranged from 0.6 to 0.8. This indicates that the Perceptual Profile is useful for the assessment of nasal fricatives.

There were differences in reliability scores for nasal fricatives in the two studies. In study 2 (repeated audio recorded assessments) there was 92% to 96% agreement and kappa scores ranged from 0.6 to 0.8. In study 1 (live analysis versus audio recorded analysis) there was 83% to 88% agreement and kappa scores from 0.3 to 0.4. This would indicate that for assessment of nasal fricatives there is a difference between live assessment and assessment based on audio recordings. Any loss of visual information in the assessment based on audio recordings was uniform when the two taped sessions were compared; however, the visual information may have influenced ratings during the live assessment. Habituation may also have been a factor in the improved reliability for the audiorecorded study.

Nasal turbulence

Reliability of assessment of strength and phoneme specificity in nasal turbulence were varied (good agreement of 83%, but poor kappa scores of 0.4). However, detailed analysis of the spread of the two ratings indicated that ten out of twelve ratings were agreed exactly, while two differed by only one point (Table 4.5). The expected agreement for strength of nasal turbulence was high, due to the fact that only two points on the scale were used in both ratings and therefore the spread of ratings was small. The high expected agreement resulted in a low kappa score.

Table 4.5 Agreement of rating of nasal turbulence strength.

	NT strength 2	
NT strength	**0**	**1**
0	1	1
1	1	9

Disagreements in ratings of nasal turbulence may be explained by the fact that in two cases weak infrequent nasal turbulence was detected in one analysis and not in another. For one of these participants, weak nasal turbulence was perceived as nasal emission. As in the previous study, the presence of a weak nasal airflow problem led to miscategorisation of the airflow errors.

Phoneme specificity in the parameter of nasal turbulence refers to the presence of nasal turbulence on one or two specific sounds. However, phoneme specificity is usually associated with nasal fricatives or velopharyngeal fricatives, which are phoneme substitutions. It was included in all four categories of nasal airflow errors in the profile in order to evaluate its position within a framework of nasality and nasal airflow errors. Phoneme specificity is not usually associated with nasal turbulence, as defined in the present study, and the poor reliability may reflect its limited value in this category.

Results indicated similar percent of agreements in study 2 and study 1 for ratings of nasal turbulence. However, the agreement for frequency of nasal turbulence improved from 79% in the live/audio study to 92% in the repeated audio recorded assessment study, and the kappa scores improved from 0.5 in study 1 to 0.8 in study 2. When strength and consistency were similar in both studies, it was surprising that frequency should improve in the second study. It may be that by removing the variable of different listening conditions, the assessment of frequency was more reliable. Or again, familiarity with the data may have influenced results. Strength and phoneme specificity of nasal turbulence had weak reliability in the two studies. This highlights two sections of the Perceptual Profile that may need to be adapted, depending on inter-rater and instrumental results.

Velopharyngeal fricative

Excellent reliability was found for strength of velopharyngeal fricatives (100% agreement and kappa score of 1). Although moderate kappa scores of 0.5 and 0.6 were found for frequency and consistency respectively, a good to excellent percent of agreement (88% and 92%, respectively) was found. The small number of occurrences of velopharyngeal fricatives may explain the discrepancy between the kappa scores and the percent agreement.

Results indicated improved reliability in study 2 for ratings of velopharyngeal fricatives. Agreement improved from acceptable (75% to 83%) in study 1 to excellent (88% to 100%) in study 2. The kappa scores also improved, with kappa scores ranging from poor (-0.1 to 0.4) in study 1 to moderate - excellent (0.5 to 1) in study 2. It is possible that the audio tape recorded speech amplified the turbulent sound during a velopharyngeal fricative, therefore reducing confusion between the perception of a nasal fricative and a velopharyngeal fricative.

4. 5. 4 Summary

This study indicated that there was good to excellent test-retest reliability of the Perceptual Profile under similar listening conditions. Although the present results are promising, they need to be considered with caution due to some limitations of the study which are highlighted in Section 4. 7.

Results indicated little difference in reliability of the assessment of nasality, nasal emission under different listening conditions. This supported a previous study (Moller & Starr, 1984). However, differences in reliability of ratings for nasal fricatives, nasal turbulence and velopharyngeal fricatives were found in the present study. These speech errors were not evaluated in the Moller and Starr (1984) study. It would appear that by using audio taped analyses, there was less confusion between the perception of nasal fricatives and velopharyngeal fricatives. The influence of habituation needs to be considered in the comparison of live/audio recorded analyses and audio/audio recorded analyses, as the rater was more familiar with the data by the end of the second study.

4. 6 Study 3: Reliability of Intra-rater Ratings using Anchor Stimuli

The anchor study was carried out five months after study 2.

4. 6. 1 Methodology

Speech Sample and Participants

The anchor study was carried out using the same audio taped speech samples as those used in the first two intra-rater reliability studies. Due to damage of one audio tape, only 8 participants were included in the anchor study.

Procedures

Before listening to the audio tapes of each participant in the study group, the rater listened to a short sample of normal speech produced by a child of similar age and the same sex (i.e., anchor stimuli). Gerratt et al. (1993) suggested that this provides the rater with a standard reference against which he/she can compare disordered speech. From the audio tapes of the study group, the rater analysed the speech samples using the Perceptual Profile. Ten weeks later another therapist not involved in the study arranged the order of the data presentation. The rater listened to the audio recording of the study group and the anchor stimuli, and rated the audio taped speech samples again. The two ratings were then compared using percent agreement and kappa analysis.

4. 6. 2 Results

Good reliability scores were found for ratings of hypernasality. Good to excellent reliability was found for ratings of nasal emission (consistency and frequency), nasal turbulence and nasal fricatives (Table 4.6). Weak reliability was found for consistency of nasality, nasal emission - strength, and velopharyngeal fricatives.

Table 4.6 Intra-rater reliability scores using anchor stimuli. * indicates too few ratings to calculate a kappa score.

Speech Parameter	% Agreement	% Expected Agreement	Kappa
Nasality	81	60	0.5
Consistency	69	66	0.09
Nasal Emission			
strength	87	78	0.4
consistency	94	73	0.8
frequency	94	73	0.8
phoneme specific*			
Nasal Fricative			
strength	87	62	0.7
consistency	94	67	0.8
frequency	87	62	0.7
phoneme specific	87	56	0.7
Nasal Turbulence			
strength	88	64	0.6
consistency	100	51	1
frequency	94	51	0.9
phoneme specific*			
Velopharyngeal Fricative			
strength	87	78	0.4
consistency	87	78	0.4
frequency	81	71	0.3
phoneme specific	75	59	0.4

4. 6. 3 Discussion

Reliability results from study 3 (anchor stimuli study) were compared to reliability results from study 2 (intra-rater reliability) in order to ascertain if there was an improvement in reliability of ratings when anchor stimuli were used. It should be noted that there were fewer number of participants in study 3 compared to study 2. This reduction in numbers may have influenced agreements in all speech categories.

Nasality

Surprisingly, reliability was weaker for nasality ratings when anchor stimuli were used. In study 3 there was 81% agreement and a kappa score of 0.5. However in study 2, when no anchor stimuli were used, there was 88% agreement and a kappa score of 0.7. The difference in percent of agreement was small, but the difference in kappa score was 0.2. This greater difference in kappa scores may be due to reduced number of participants in study 3. Another important factor in evaluating the results of the two studies may be habituation and the time lag between studies. Although the rater had analysed the data three times previously, there was a time lag of five months between study 2 and study 3. This may have resulted in less familiarity with the data in study 3 compared to study 2.

The present results using anchor stimuli compare well with a previous study by Watterson et al. (1993). They reported good agreement (Pearson product moment correlation coefficient of 0.88) for intra-rater reliability using anchor stimuli during the practice session only, but unlike the present study, the anchor stimuli was not used prior to the analysis of each speech sample. In the present study, the lack of substantial improvement in study 3 using anchor stimuli questions the benefit of anchor stimuli in the assessment of nasality.

Consistency of nasality showed almost no agreement in the present anchor stimuli study. This may be explained by the discrepancies in ratings of nasality. In cases where there was discrepancy in nasality ratings, the consistency was also rated differently; for example, one participant had a nasality rating of 'd' (moderate to severe) and was also rated inconsistent (i.e. this participant was occasionally rated as 'c' (moderate)). During the second analysis the participant's nasality was rated as 'c' and consistent. It seems that the moderate to severe (d) rating of nasality was not detected in the second analysis; therefore, the rating of consistency was different in each analysis. This scenario was evident in four of the cases studied.

Nasal emission

There was little difference in reliability scores for nasal emission in the studies with and without anchor stimuli. In study 3, there was 87% agreement and a kappa score of 0.4 for strength of nasal emission. In study 2, where no anchor stimuli were used, there was

91% agreement and a kappa scores of 0.6. This indicated marginal improvement when anchor stimuli were not used. Consistency of nasal emission had 94% agreement and a kappa score of 0.8 in the present study compared to 91% agreement and a kappa score of 0.5 in study 2. The higher kappa score in study 3 may be explained by the reduced number of participants in that study. There was limited difference in frequency of nasal emission, with 94% agreement and a kappa score of 0.8 in study 3 compared to 92% agreement and the same kappa score in study 2. Results indicated no benefit in using anchor stimuli for the assessment of nasal emission.

Nasal fricative

Reliability scores for a nasal fricative was similar in both studies. In study 3, using anchor stimuli, there was 87% agreement and a kappa score of 0.7 for strength of nasal fricative, compared to 92% agreement and a kappa score of 0.7 in study 2. Consistency of nasal fricative had 94% agreement and a kappa score of 0.8 in study 3, and 96% agreement and a kappa score of 0.8 in study 2. For frequency of nasal fricative there was 87% agreement and a kappa score of 0.7 in study 3 compared to 96% agreement and a kappa score of 0.8 in study 2. These results indicated an improved percentage of agreement when anchor stimuli were not used.

Nasal turbulence

Consistency of nasal turbulence had improved reliability using anchor stimuli. In the anchor stimuli study there was 100% agreement and a kappa score of 1, while in the study without anchor stimuli there was 83% agreement and a kappa score of 0.6. There was minimal improvement for reliability of strength of nasal turbulence when anchor stimuli were used (88% agreement and a kappa score of 0.6 with anchor stimuli, 83% agreement and a kappa score of 0.4 without anchor stimuli). There was no difference in reliability for frequency of nasal turbulence.

Velopharyngeal fricative

In general there were weaker reliability scores for velopharyngeal fricatives when anchor stimuli were used. In the anchor stimuli study there was 87% agreement for all parameters of velopharyngeal fricatives. In study 2, where no anchor stimuli were used, the percentage of agreement ranged from 88% to 100%. Kappa scores in study 2 were greater than kappa scores in study 3 despite the fact that smaller numbers of participants were used in study three.

4. 6. 4 Summary

The results of studies 2 and 3 indicated minimal differences in reliability results using the Perceptual Profile when anchor stimuli were used. Results indicated marginally better reliability scores for nasal airflow errors in the study where anchor stimuli were not used. The only improvement using anchor stimuli was found to be in the reliability of consistency of nasal turbulence. One can conclude that under the conditions tested, the use of anchor stimuli did not improve the reliability of the Perceptual Profile. Factors such as habituation and differences in the time lag between studies need to be considered. However, if results from study 1 are compared to results from study 3, where habituation should have been reduced due to the time lag of seven months, there is no overall improvement when anchor stimuli are used. In fact, reliability for nasality and nasal emission was better without anchor stimuli. The percent agreement for nasal fricatives was the same in the two studies, while percent agreement and kappa were lower without anchor stimuli. Results from the above studies are inconclusive in evaluating the use of anchor stimuli.

One possible reason for the lack of improvement using anchor stimuli may be the manner in which the anchor stimuli were used. In a previous study by Gerratt et al. (1993) anchor stimuli were used for assessment of voice quality. The stimuli that were presented to the listener represented varying degrees of abnormal voice quality against which the listener compared the speech sample to be rated. In the present study, normal speech was used as an anchor stimulus. This was in an attempt to provide the listener with a baseline for normal nasality and nasal airflow errors. Another methodological difference between the present study and previous studies was that the anchor stimuli were presented prior to rating the speech sample. However, in the study by Watterson et al. (1993) the anchor stimuli were used during the training session only and not immediately prior to rating the speech sample. The different uses of anchor stimuli requires further investigation in order to evaluate the value of anchor stimuli in improving reliability of ratings.

4. 7 Limitations of the Intra-rater Reliability Studies

The author who also devised the Perceptual Profile, due to the time constraints within the clinical setting, carried out intra-rater reliability studies. Not only was the author was highly familiar with the Perceptual Profile, but she also had many years experience working in the area. This is supported by some of the results where the agreement was consistently greater than the expected agreement calculated for the kappa scores. As a result, the intra-rater reliability may have been biased and can only be applied to one rater. Hence, intra-rater reliability results are difficult to generalise.

It was considered important to keep the speech samples consistent across all reliability studies. One advantage of this is that it reduces the variably between studies. However, one disadvantage is the possibility of habituation of the listener. In the present study, the same speech samples were analysed on five different occasions, and this may have resulted in familiarity with the data. The slightly better results found in study 2 may be attributed to some extent to familiarity with the speech data, since the second analysis took place relatively soon after the live analysis. The variation in time between the initial and second analyses may also have influenced results. For some of the data in study 1, there was only a three week gap between analyses; however, in study 2 and 3, there was a ten week gap. This longer gap was intended to reduce habituation. Another variation in time was the interval between each study. The five month time lag between study 2 and 3 may have reduced the degree of familiarity with data.

A further limitation of the present study was that there were a different number of participants in study 2 and study 3 due to damage of one audio tape. One may have expected greater percentage of agreement in study 3 due to the smaller number of participants in the study; however, this was not the case.

In general the audio taped speech samples were analysed having listened once to the tape. However, occasionally, the tape was played a second time if the rater was undecided about a rating, and again this may have influenced results. The importance of visual clues such as nasal and facial grimace has been highlighted in the literature (McWilliams and Phillips, 1979; McWilliams et al., 1990). One would expect improved reliability from live assessments; however, this is impossible to test due to individual variation in speech production from one situation to another (Laver, 1980; Kuehn, 1982).

Only 12 participants were included in the intra-rater reliability studies. Because of the detail of the Perceptual Profile, there were a limited number of occurrences of all the possible errors in the intra-rater reliability studies. As a result the kappa scores may have underestimated the reliability between the two ratings. Further investigations using a larger population are required.

4. 8 Intra-rater Reliability Studies - Conclusions

The three intra-rater reliability studies indicated good reliability for the rater in the present study using the Perceptual Profile. This indicates that a rater who is familiar with the profile and who has experience in the area of nasality and nasal airflow errors in speech can reliably assess speech errors using the profile. The studies indicate that the perceptual assessments of nasality and nasal airflow errors, which will be used in the instrumental study (Chapter 5), are reliable. Unfortunately, these results cannot be generalised to other raters due to some of the limitations of the present reliability studies. It is recommended that further intra-rater reliability studies be carried out, following inter-rater reliability analyses and comparison of the perceptual and instrumental assessments.

Results indicated minimal differences between listening conditions for assessment of nasality, nasal emission and nasal turbulence. However, differences in reliability between repeated audio taped analyses and live/audio analyses were found for assessment of nasal fricatives and velopharyngeal fricatives. Results indicated that when there was a difference in listening conditions, reliability was weaker. This may be because visual information is lost during audio recordings. It may also result from increased familiarity with the data in the audio/audio recorded analyses. However, in view of the importance of visual information in assessment of nasal airflow errors, it was decided to compare live perceptual assessments of nasality and nasal airflow errors with the instrumental assessments in the main study.

Results indicated no improvement in reliability of the Perceptual Profile when anchor stimuli were used. In comparing the results of study 1 and study 3, and the results of study 2 and study 3, there was no advantage in using the anchor stimuli. Due to the lack of significant improvement in reliability scores using anchor stimuli, the use of anchor stimuli in the main study was not indicated.

Comparison of the three intra-rater reliability studies indicated that the weakest reliability scores were found for ratings of velopharyngeal fricatives. This questions the feasibility of using a separate speech category for velopharyngeal fricatives on the Perceptual Profile. In all three studies disparity of ratings could be explained by weak airflow errors. These issues will be discussed in further detail below in the content of inter-rater reliability results.

4. 9 Study 4: Inter-rater Reliability Study

4. 9. 1 Methodology

Participants

Audio taped speech samples of twenty children, aged between 3;4 years and 13;10 years, were included in the inter-rater study. Of these twenty participants, ten had been included in the intra-rater studies and ten had been included in the pilot study (Appendix 12). Children with cleft palate, velopharyngeal dysfunction, or nasal obstruction were included. All participants presented with abnormal nasality and/or nasal airflow problems, with or without articulatory errors. Children were excluded if there was evidence of any of the following: severe dyspraxia/dysarthria; learning disability (greater than mild); bilateral hearing loss above 45 db.; nasal blockage associated with an upper respiratory infection; moderate to severe hoarseness of voice; mixed nasal resonance; and an inability to complete the assessment protocol due to behavioural problems or poor co-operation.

Raters

Three raters were included in the study. One rater (TS) was a specialist speech and language therapist in cleft palate and velopharyngeal dysfunction, with nineteen years experience. The second rater (EM) had worked for nine years part-time as a speech and language therapist on the cleft palate team and the third rater (MO'M) had limited experience in this area.

Training

Each of the two raters (EM and MO'M) had an individual two hour training session given by TS prior to the study. The training sessions were carried out separately due to time constraints within the clinic. Training included:

- definitions of each category on the Perceptual Profile with audio taped demonstrations;
- definitions of weak vs. strong, frequent vs. infrequent, consistent vs. inconsistent, and

phoneme specific with audio taped demonstrations;

- the use of anchor stimuli was explained and demonstrated;
- four samples of speech were analysed by the trainer and trainee simultaneously from audio tape recordings, the results were discussed and a consensus for the ratings was reached.

Procedures

The same speech sample, and recording procedure outlined above for the intra-rater study 1 was used. Raters rated audio taped speech samples which had been recorded on the Sony Scoopman digital audio tape recorder using a Sony condenser microphone. The microphone had been placed at the level of the child's mouth, approximately four inches away. A recording level between 25db and 28db was used. Audio playback was carried out in a quiet room at a decibel level of 20db to 25db, with the use of earphones. Prior to judging each participant the rater listened to the anchor stimuli. The rater then listened to the test speech sample and classified each of the following levels of speech - words, sentences, automatic speech and conversational speech according to the perceptual framework (see section 5.2). When the speech sample had been rated, the rater made a final judgement and recorded it in the top left hand section of the assessment sheet (Appendix 13b). The most severe degree of nasality and/or nasal airflow disorder based on all speech samples was recorded.

4. 9. 2 Inter-rater Reliability Results

Paired inter-rater reliability was analysed for 3 raters (TS, MO'M and EM), yielding percent agreement and kappa scores for each pair of raters. Inter-rater reliability between TS and MO'M was based on analysis of 20 cases; however, due to damage of one tape, reliability between EM and the other two raters was based on 18 cases. Agreement and kappa scores for the three pairs of raters are presented for each speech parameter separately.

Nasality

Inter-rater reliability for ratings of nasality ranged from moderate to good. Percentage of agreement ranged from 77% to 85%, and kappa scores ranged from 0.4 to 0.6 (Table 4.7). Consistency of nasality had weaker reliability ranging from poor to good, with agreement ranging from 61% to 85%, and kappa scores ranging from 0.1 to 0.7.

Table 4.7 Inter-rater reliability (percent agreement and kappa scores) of ratings of nasality and consistency of nasality for three pairs of raters.

	TS/MO'M		TS/EM		EM/MO'M	
Speech Parameter	**Percent agreement**	**Kappa score**	**Percent agreement**	**Kappa score**	**Percent agreement**	**Kappa score**
Nasality	85	0.6	79	0.4	77	0.4
Consistency	61	0.1	69	0.3	85	0.7

Nasal emission

Inter-rater reliability for ratings of nasal emission varied considerably between pairs of raters. Percentage of agreement ranged from 75% to 86%, and kappa scores ranged from 0.2 to 0.5 (Table 4.8).

Table 4.8. Inter-rater reliability (percent agreement and kappa scores) of ratings of nasal emission for three pairs of raters.

	TS/MO'M		TS/EM		EM/MO'M	
Speech Parameter-nasal emission	**Percent agreement**	**Kappa score**	**Percent agreement**	**Kappa score**	**Percent agreement**	**Kappa score**
Strength	75	0.2	78	0.2	78	0.3
Consistency	77	0.3	86	0.2	78	0.3
Frequency	77	0.3	86	0.2	78	0.3
Phoneme specific	70	0.2	78	0.2	78	0.5

Nasal fricative

Reliability for ratings of nasal fricatives varied considerably. The percent agreements were good to excellent (83% to 92%), while kappa scores ranged from poor to good (-0.07 to 0.6) (Table 4.9).

Table 4.9 Inter-rater reliability (percent agreement and kappa scores) of ratings of nasal fricatives for three pairs of raters.

	TS/MO'M		TS/EM		EM/MO'M	
Speech Parameter- Nasal fricative	**Percent agreement**	**Kappa score**	**Percent agreement**	**Kappa score**	**Percent agreement**	**Kappa score**
strength	90	0.6	86	-0.07	86	0.4
consistency	90	0.6	86	-0.07	86	0.4
frequency	92	0.6	86	-0.07	89	0.5
phoneme specific	92	0.5	83	-0.08	89	0.3

Nasal turbulence

Inter-rater reliability for ratings of nasal turbulence ranged from poor to good. The percentage of agreement ranged from 72% to 83% and kappa scores ranged from 0.3 to 0.5 (Table 4.10).

Table 4.10 Inter-rater reliability (percent agreement and kappa scores) of ratings of nasal turbulence for three pairs of raters.

	TS/MO'M		TS/EM		EM/MO'M	
Speech Parameter-Nasal turbulence	**Percent agreement**	**Kappa score**	**Percent agreement**	**Kappa score**	**Percent agreement**	**Kappa score**
strength	80	0.3	83	0.3	83	0.5
consistency	80	0.5	72	0.4	72	0.4
frequency	76	0.4	72	0.4	78	0.5
phoneme specific	80	0.4	78	0.4	78	0.5

Velopharyngeal fricative

Inter-rater reliability for ratings of velopharyngeal fricatives varied. The percent agreement ranged from poor to good (69% to 86%) and the kappa scores ranged from poor to moderate (0 - 0.5) (Table 4.11).

Table 4.11 Inter-rater reliability (percent agreement and kappa scores) of ratings of velopharyngeal fricative for three pairs of raters.

	TS/MO'M		TS/EM		EM/MO'M	
Speech Parameter-nasal emission	**Percent agreement**	**Kappa score**	**Percent agreement**	**Kappa score**	**Percent agreement**	**Kappa score**
strength	82	0	80	0.5	69	0
consistency	77	0	81	0.5	72	0
frequency	75	0	83	0.5	78	0
phoneme specific	82	0	86	0.5	80	0

Multi-rater kappa scores

In order to ascertain agreement between all three raters on the Perceptual Profile, multi-rater kappa scores were calculated for all speech parameters (Fleiss, 1981). Multi-rater kappa analysis was calculated for 18 participants (Table 4.12). Results indicated poor to good kappa scores on the speech parameters.

Table 4.12 Multi-rater kappa scores for three raters for ratings of all speech parameters.

Speech Parameter	Multi-rater kappa scores
Nasality	0.3
Consistency	0.3
Nasal Emission	
strength	0.1
consistency	0.
frequency	0.
phoneme specific	0.3
Nasal Fricative	
strength	0.6
consistency	0.4
frequency	0.4
phoneme specific	0.3
Nasal Turbulence	
strength	0.3
consistency	0.35
frequency	0.3
phoneme specific	0.4
Velopharyngeal Fricative	
strength	0.08
consistency	0.1
frequency	0.2
phoneme specific	0.2

4. 9. 3 Inter-rater Reliability - Discussion

Nasality

Reliability of ratings of nasality between TS and MO'M was good (85% agreement and a kappa score 0.6). However, reliability between EM and the other two raters was weaker (77/79% agreement and a kappa score of 0.4). Previous studies have indicated varying reliability results for inter-rater ratings of nasality, using different statistical analyses of reliability (Table 4.13). Comparison with previous studies is difficult because of the different types of analyses used. The percent agreement found in the present study compares well to previous studies (Sweeney, 1984; Lohmander-Agerskov, 1996; Watterson et al., 1998a). Using previous guidelines (Lohmander-Agerskov, 1996; Watterson et al., 1998b), we can conclude that the percent agreement for ratings of nasality is acceptable. The kappa scores in the present study did not compare well with the kappa scores found in the CSAG (1998) study. However, the scale used in the CSAG (1998) study was considerably shorter and there was a longer training period for the raters. The present study indicates that the inter-rater reliability for ratings of nasality using the Perceptual Profile is inconclusive and that further investigation is recommended.

Table 4. 13 Inter-rater reliability results from previous studies and the present study.

Study	Reliability Analysis	Scale	Result
Sweeney (1984)	Percent agreement	6 points	77% - 87%
Sell & Grunwell (1990)	Spearman Rank correlation	5 points	0.74
Nellis et al. (1992)	Intraclass correlation	6 point	0.91
Hardin et al. (1992)	Pearson correlation	5 points	0.82
Lohmander-Agerskov (1996)	Percent Agreement	5 points	70% - 80%
Watterson et al. (1998a)	Percent agreement	5 points	80%
CSAG (1998)	Kappa score	3 points	0.8
Pinborough-Zimmerman (1999)	Intraclass correlation	6 points	0.91
Present Study	Percent agreement Kappa score	5 points	77% - 85% 0.4 - 0.6

Analysis of the previous results raises an important question regarding the reliability results reported in these studies. As Watterson et al. (1998b) pointed out, percent agreement and correlations do not always indicate reliability. The reliability of previous perceptual scales is therefore questionable. The present study highlights the importance of rigorous reliability testing using appropriate statistics.

The multi-rater kappa score for ratings of nasality was 0.3; however, this score may have been poor because of the weaker agreement between EM and the other two raters. No previous study reported multi-rater kappa scores.

Agreement on ratings of consistency of nasality varied considerably between raters. Poor reliability (agreement of 61% and a kappa score 0.1) was found between TS and M'OM, whereas good reliability (agreement of 85% and a kappa score of 0.7) was found between EM and MO'M. This discrepancy is difficult to explain. It is possible that the poor agreement for the rating of consistency of nasality between TS and MO'M may be due to unclear definitions of consistency and hence, different interpretations of consistency by the two raters.

Nasal emission

Agreement between raters for ratings of nasal emission ranged from 70% to 86%, the weakest agreement was between phoneme specificity (70% to 78%) and strength of nasal emission (75% to 78%). Good agreement was found between consistency and frequency of nasal emission (77% to 86%). The present results compare favourably with two previous studies. Lohmander-Agerskov (1996) reported 70% to 80% agreement using a 5 point scale. An agreement between the presence and absence of nasal emission of 77% was reported by Sell et al. (1990). In the latter study, there was no attempt to measure severity or consistency. The nasal airflow errors were not differentiated into categories of nasal emission, nasal turbulence, nasal fricatives or velopharyngeal fricatives. Despite the greater detail in the present study results indicated similar and improved levels of agreement.

Kappa scores for ratings of nasal emission in the present study were weaker than kappa scores reported in the CSAG (1998) study. In the present study, kappa scores ranged from 0.3 to 0.5; however, in the CSAG study a kappa score of 0.6 was reported using a short rating scale. In the present study there were effectively 8 possible ratings for nasal emission. One explanation for poor reliability using kappa analyses was that there was a small number of occurrences of nasal emission in this study. The kappa score was calculated by comparing expected agreement (chance agreement) with observed agreement. The expected agreement was calculated from the number of occurrences of the speech error. If the error had a large number of 0 ratings (i.e. absent), then the expected agreement was high. When a high expected agreement was compared to a high observed agreement, the kappa score was low.

Discrepancies in ratings for nasal emission may be due to the loss of visual information as assessments were made from audio recorded speech samples. McWilliams and Phillips (1979) report that it is difficult to assess nasal emission reliably from audio tape recordings, as the visual clues of nasal emission are not evident. If the facial or nasal grimace that is often associated with nasal emission cannot be seen, there may be discrepancies in ratings. The discrepancies in ratings for nasal emission may also have been due to confusion between nasal turbulence and mild nasal emission. Nasal turbulence was perceived as the primary problem in three participants; however, one rater also recorded mild nasal emission in these participants. It is possible that this mild nasal emission was rated as nasal turbulence by the other two raters. If this was the case, it may be that listeners rated nasal emission and nasal turbulence as a continuum and not as separate error categories, which lends support to the concept that nasal turbulence is a more severe form of nasal emission as described by Sell et al. (1994).

Nasal fricative

No previous study reported reliability for ratings of nasal fricatives. In the present study, the percent of agreement of ratings for nasal fricatives was good for all three pairs of raters (83% to 92%). There was a large variation in kappa scores. Kappa scores of 0.5 to 0.6 were found for TS and MO'M; however, the kappa scores for TS and EM were low and negative (-0.07). Again, the discrepancy noted between percent agreement and kappa scores may be due to the large number of 0 ratings by both raters. This resulted in a high percentage score but a low kappa score. The spread of ratings for 18 participants is presented to illustrate the high incidence of 0 ratings in this category (Table 4.14)

Table 4.14 Ratings for nasal fricative - strength by TS and EM.
0 indicated absence of nasal fricative, 1 indicates weak nasal fricative and 2 indicates strong nasal fricative.

NF strength (EM)

NF strength (TS)	0	0	2
0	15	0	1
1	1	0	0
2	1	0	0

Results indicated that although TS and EM had 86% agreement on ratings for nasal fricative - strength, the kappa score was -0.07. However, EM and MO'M also had 86% agreement on ratings of nasal fricative - strength, but the kappa score was 0.4. This discrepancy in kappa scores may be explained by the higher 'expected agreement' (87%) based on the spread of ratings between TS and EM compared to the 'expected agreement' (77%) between EM and MO'M. The table of ratings by EM and MO'M illustrates the spread of ratings for the two raters (Table 4.15). Comparing Table 12 and 13 highlights the variation in spread of ratings for the two pairs of raters. Because of the small number of occurrences and the small number of participants, slight variation in the spread of ratings resulted in large discrepancy in kappa scores.

Table 4.15 Ratings for nasal fricative - strength by MO'M and EM.
0 indicates absence of nasal fricative, 1 indicates weak nasal fricative and 2 indicates strong nasal fricative.

NF strength (EM)

NF strength (MO'M)	0	0	2
0	14	0	0
1	1	0	0
2	2	0	1

Interestingly, the multi-rater kappa scores for nasal fricatives - strength were good (0.6). Overall results indicated moderate reliability on ratings of nasal fricatives.

Nasal turbulence

No previous study reported reliability for inter-rater ratings of nasal turbulence. In the present study, agreement for ratings of nasal turbulence ranged from 72% to 83%, with good agreement for all pairs on ratings of nasal turbulence - strength (83%). Kappa scores ranged from poor to moderate (0.3 to 0.5). As with nasal fricatives, there were incidences where there were similar percents of agreement but different kappa scores. The difference in kappa may have been due to the distribution of ratings between pairs. For example, TS had ratings of 0 to 1 (Table 4.16), whereas EM and MO'M had ratings of 0 to 2 (Table 4.17). However, although kappa scores were low, ratings never differed by more than one point on the scale.

Table 4.16 Ratings for nasal turbulence - strength by TS and EM.
0 indicates absence of nasal fricative, 1 indicates weak nasal fricative and 2 indicates strong nasal fricative.

NT strength (EM)

NT strength (TS)	**0**	**1**	**2**
0	2	0	0
1	3	10	3
2	0	0	0

Table 4.17 Ratings for nasal turbulence - strength by MO'M and EM.
0 indicates absence of nasal fricative, 1 indicates weak nasal fricative and 2 indicates strong nasal fricative.

NT strength (EM)

NT strength (MO'M)	**0**	**1**	**2**
0	3	2	0
1	2	7	1
2	0	1	2

Inter-rater reliability results indicated moderate reliability of the Perceptual Profile in ratings of nasal turbulence.

Velopharyngeal fricative

Inter-rater reliability for ratings of velopharyngeal fricatives has not been reported previously. The agreement on ratings for velopharyngeal fricatives in this study was good for TS and EM (80% to 86%), with an acceptable kappa score of 0.5. Weaker agreements were found between MO'M and the other two raters. This was due to the fact that MO'M did not distinguish between velopharyngeal fricatives and nasal fricatives. MO'M rated all participants as 0 for velopharyngeal fricatives. The velopharyngeal fricatives rated by the other two raters were weak in most cases.

4. 9. 4 Limitations of the Study

One major drawback in terms of reliability of the Perceptual Profile was the number of descriptive points on the scale. Reliability of a scale decreases as the number of points on the scale increases (McWilliams et al., 1990; Henningsson and Hutters, 1997). On this profile, there were eighteen possible descriptions for each child's speech. Henningsson and Hutters (1997) stated that reliability may be increased by reducing the number of scale points on the scale but the information gained may not then be sufficiently detailed. If fewer number of points are used on a scale, it may be too coarse to measure relevant perceptible differences before and after treatment, and importantly, it may not be useful in evaluating the relationship with instrumental measures. The detail of the Perceptual Profile in the present study may account for weak inter-rater kappa scores. Reduction in the size of the scale should improve reliability of the profile, but adequate information must be maintained.

Insufficient training in the descriptive categories and use of the Perceptual Profile is probably a significant limitation of the present study. The two raters had been familiar with the profile and the definitions prior to training; however, their ability to use the definitions and the Profile had been over estimated by the trainer. For example, poor reliability in consistency of nasality may have been due to unclear interpretations of the term consistency. However, ratings of consistency of nasal airflow errors were reliable.

Perhaps listeners did not use the term in the same manner for nasality and nasal airflow errors. Furthermore, it is also significant that training occurred on separate occasions for the two trainees, with possible different emphases given to the areas of training. At no time therefore was consensus listening undertaken by all three listeners at the training stage.

Raters had different degrees of experience and familiarity in the area of nasality and nasal airflow errors in speech, which may have influenced the results. The highly experienced author was highly familiar with the types of speech errors being assessed as well as with the Perceptual Profile, and was a specialist in the area. TS was also more familiar with the data as it had been used in the previous intra-rater study and the pilot study. EM had 9 years experience in the area, but was not considered a specialist as she had received no post graduate training in the area). Furthermore, EM had been on six months leave from her employment at the time of the study and therefore may have had reduced familiarity with the speech errors and the Profile. Interestingly, higher overall scores were found between TS and MO'M (the least experienced listener), who had limited experience in the area of nasality and nasal airflow errors in speech. This may be due to the fact that MO'M received all her training in this area from TS. Finally, only one of the listeners was a specialist speech and language therapist in the area of cleft palate and velopharyngeal dysfunction.

There was variation in the number of tapes analysed in the inter-rater reliability study. Percent agreement and kappa scores between EM and the other raters were calculated on 18 cases as opposed to 20. This may have contributed to the slight differences in percentage of agreement and kappa scores found between raters.

In general, low kappa scores reflect the small number of cases presenting with each speech parameter. It has been recommended that percent of agreement and reliability should be reported when analysing ratings of speech (Kreiman et al., 1993). Because of the detail of the Perceptual Profile, analysis of a large number of speech samples would be required in order to evaluate kappa scores more accurately.

4. 9. 5 Inter-rater Reliability Study -Summary

In general the percent of agreement of inter-rater ratings of nasality and nasal airflow errors ranged from acceptable (75%) to good (90%). No other studies have reported reliability of ratings of nasal turbulence, nasal fricatives or velopharyngeal fricatives. Only one study has reported the inter-rater reliability for ratings of nasal emission (CSAG, 1998). The percent of agreement found in the present study compares well with previous studies, notwithstanding the detail of the profile. Reliability analyses using kappa scores were disappointing. Results of the kappa analyses did not indicate that the Perceptual Profile was reliable across different raters for the assessment of nasal emission, nasal turbulence and velopharyngeal fricatives.

It is possible that raters heard the same speech parameters during assessment, but classified them differently. If this was the case, then further training on definitions and classification should increase reliability. Further investigation regarding training would help identify the number of hours of training required in order to reach satisfactory levels of agreement.

This study highlights the variation between listeners when assessing children with nasality and nasal airflow problems. Varying degrees of listener experience and training in the perception of disordered speech may have contributed to the variation between listener ratings (Kreiman et al., 1990). In previous reliability studies, specialist speech and language therapists in cleft palate and velopharyngeal dysfunction rated the speech samples (Sell & Grunwell, 1990; Watterson et al., 1993). In the present study, only one rater was a specialist, while the other two raters each had very different levels of experience in the area. It is recommended that further inter-rater reliability studies be carried out using speech and language therapists specialising in cleft palate and velopharyngeal dysfunction.

4.9. 6 Inter-rater Reliability - Conclusions

The descriptive scale used in the Perceptual Profile is a novel type of assessment approach and the inter-rater agreements are encouraging. However, the variable kappa scores indicate inconclusive reliability. The percent agreement found using the Perceptual Profile is similar to that found in other scales, despite its greater detail. However, when more rigorous analyses of reliability are undertaken, the overall reliability of the Profile is weak. This study raises the question of reliability of any assessment of nasality and nasal airflow errors in speech, since previous studies have not undertaken such rigorous analyses. Further investigations into the reliability of the Perceptual Profile are recommended with considerations of training, reduction in detail of the Profile and possibly the validity of the Profile.

4. 10 General Discussion of Reliability Results

Recent literature stressed the importance of calculating both the percent of agreement between ratings and a statistical agreement that incorporates the influence of chance agreement (kappa values). In the present study these agreements were calculated; however, results varied according to the type of analysis (i.e. percent of agreement or kappa scores). Where both the percent of agreement and the kappa scores were good, one could have confidence in the reliability of the assessment tool. Hence, there was good intra-rater reliability for ratings of nasality and nasal emission. Intra-rater reliability for ratings of nasal turbulence was acceptable, however there was questionable intra-rater reliability for ratings of nasal fricatives and velopharyngeal fricatives (where there was an acceptable percent of agreement but lower kappa score). Inter-rater reliability varied considerably between listeners. In general, there was good reliability for ratings of nasality and nasal fricatives. Reliability on other airflow errors was questionable. Where reliability scores varied according to the type of analysis, one needs to question why there was an adequate percent of agreement, but low kappa scores.

Discrepancies in ratings of nasal airflow errors were more prevalent when the nasal airflow error was mild. It is possible that, perceptually, there is an overlap between weak nasal emission and weak nasal turbulence in terms of velopharyngeal function and respiratory effort. Kummer et al. (1992) reported that participants with nasal emission had greater velopharyngeal gap size during sound production than participants with nasal turbulence. When nasal turbulence is present the velopharyngeal gap is sufficiently narrow to produce friction. It is possible that when nasal emission is weak or mild the velopharyngeal gap size is smaller than when nasal emission is strong or severe. Hence, nasal emission and nasal turbulence may exist as a continuum reflecting differences in velopharyngeal closure. This hypothesis requires further investigation using the perceptual assessment and nasendoscopic assessment of velopharyngeal function. Results indicated that weak nasal airflow errors led to confusion in categorisation of errors. This profile is apparently less reliable at the mild end of the scale compared to the moderate or severe end of the scale.

In these reliability studies, there was confusion between ratings of nasal fricatives and velopharyngeal fricatives. In previous reports by Grunwell and Harding (1996) and Harding and Grunwell (1998), no distinction was made between nasal fricatives and velopharyngeal fricatives. Grunwell and Harding (1996) stated that a nasal fricative involved complete oral closure, with noisy or turbulent nasal airflow. Although the present definitions distinguished between nasal fricatives and velopharyngeal fricatives, there was some difficulty perceiving these differences. It is possible that as with nasal emission and turbulence, a weak velopharyngeal fricative was perceived as a nasal fricative. Consideration should be given combining nasal fricatives and velopharyngeal fricatives into a single category. However, this distinction will be required at this stage in order to examine the relationship between perceptual instrumental assessments. This will be discussed further in Chapter 9, informed by the results of the instrumental investigation.

4. 11 Conclusions from Reliability Studies

Results from the present study indicate that there is good intra-rater reliability for the author of the study using the Perceptual Profile. The Perceptual Profile has adequate intra-rater reliability for comparison with instrumental measurements in the main study. However, these intra-rater reliability results cannot be generalised to other raters, and further intra-rater reliability studies are recommended. Inter-rater reliability results were inconclusive. There were acceptable levels of agreement between ratings of nasality and nasal airflow errors, however kappa scores were poor. Where it was possible to compare the reliability of the Perceptual Profile with previous studies, the Perceptual Profile compared favourably. As a result, it can be considered a useful assessment tool. However, the profile will require adaptation in order to improve reliability. Reduction of the categories of nasal fricative and velopharyngeal fricative into one category may be useful. Final adaptations will consider the present reliability results and results of comparison between perceptual assessments and instrumental measurements. The present inter-rater reliability study has emphasised the need for further training in the use of the Profile, reduction in detail of the Profile, and an increase in the sample size to include all nasal airflow errors. It also raises the question of reliability in previous studies of nasality and nasal emission where rigorous statistical analyses were not carried out.

4. 12 Implications of Results for the Instrumental Study

Live perceptual analyses will be used to compare perceptual measurements of nasality and nasal airflow errors with instrumental measurements of nasal acoustic energy and pressure/flow. The Perceptual Profile has been found to have good reproducibility, and variable reliability. However the importance of visual clues, such as nasal and facial grimace has been underscored (Sell et al., 1994; Henningsson and Hutters, 1997). As a result of this and the good intra-rater reliability between live and audio taped analyses, it was decided to use live analyses for the instrumental study. The audio recordings may be used for further analyses, if required. Anchor stimuli will not be used in the instrumental study as no improvement in reliability was found using anchor stimuli.

CHAPTER 5

METHODOLOGY

Specific criteria were identified for the selection of assessment tools for the present study. The assessment tools should to be easy to use and calibrate (assessments were undertaken in a busy out patient clinic where lengthy calibration, collection and analysis of data were impossible). The time required to complete an assessment should be sufficiently short in order to prevent fatigue. The assessment technique should be non invasive so that it was possible to assess children from 4 years upwards; accurate and reproducible. The computerized data should be stored for later analysis. In addition normative data should be available for the instrumental assessment and the instrumental equipment should be commercially available and affordable.

The Perceptual Profile was developed for this study with these criteria in mind. Results of the reliability study indicated good intra-rater reliability for the examiner in this study. The Nasometer is extensively used, both clinically and in research to assess nasal resonance. It is easy to use and calibrate, non invasive, and assessment was short. Previous studies reported normative data for various language, dialects and speech stimuli (Fletcher et al., 1989; Seaver et al., 1991; Haapanen, 1992; Watterson et al., 1996; Trindade et al., 1997; van Doorn & Purcell 1998) (see Chapter 3. 4). Reliability and validity have been reported in the literature (Nichols, 1999; Watterson et al., 1999). PERCI-SARS was selected for the pressure/flow measurements as it is easy to use and calibrate, suitable for children from 4 years upwards, and had a good response time (1 millisecond). As part of the selection process of the aerodynamic equipment, the NORS had been tested, but was found to be unreliable and have poor validity (see Section 2. 7. 1). Investigations indicated that the Gaeltec was not commercially available and the paediatric divided Rothenberg mask was too large to be used with small children. Therefore these devices did not meet the criteria for this study. A functional specification for an airflow device was devised by the author and a Bio-medical Engineer and sent to the manufacturers of the PERCI-SARS (Appendix11). The feedback from the specification and independent reviews by two users of the system indicated that this system was the most suitable system available for the aerodynamic study.

5.1 Participants

The study group consisted of a consecutive series of children (n = 50) who were referred to a national Cleft Palate Unit for investigation of speech problems. Children with cleft palate, velopharyngeal dysfunction, or nasal obstruction, with or with out syndromes, were included. All participants presented with abnormal nasality and/or nasal airflow problems, with or without articulatory errors. Age of the participants ranged from 4;10 years to 15;10 years.

Children were excluded if there was evidence of any of the following: severe dyspraxia/dysarthria; learning disability (greater than mild); bilateral hearing loss above 45 db.; nasal blockage associated with an upper respiratory infection; moderate to severe hoarseness of voice; mixed nasal resonance; and an inability to complete the assessment protocol due to behavioural problems or poor co-operation.

Hearing acuity was assessed within one month of the speech assessment.

5. 2 Materials

5. 2. 1 Perceptual profile

A classification system to describe nasality and nasal airflow errors has been developed. The system describes nasality (hypernasality and hyponasality) and nasal airflow in an auditory perceptual framework (Chapter 4).

Nasality was assessed on a descriptive scale where the presence or absence of nasality problems were recorded. If present, the deviation from the norm was described. Hypernasality and hyponasality were described as consistent or inconsistent. The presence or absence of cul de sac resonance was recorded.

Nasal airflow is divided into 4 categories - nasal emission, nasal fricative, nasal turbulence and velopharyngeal fricative. Each category is described in an auditory perceptual framework where audibility, frequency, consistency and phoneme specificity of the airflow is recorded (Appendix 12b)

5. 2. 2 Nasometer

Nasal resonance was measured using the Nasometer model 6200.3 manufactured by Kay Elemetrics. A headset, containing a sound separator with microphones on either side, detected the oral and nasal acoustic components of the participant's speech (Fig 5.1). Each microphone signal was amplified by a preamplifier. The microphones had a bandwidth of 15,000 Hz and were balanced for equal gain as part of the calibration. The signals were filtered with a 300-Hz band-pass filter that had a centre frequency of 500 Hz. This ensured that the selected frequencies of the oral and nasal signals were captured, while the high frequency and low frequency ambient noise were rejected. The acoustic signals were then fed to an analogue to digital converter (response time of 35 ms). The data acquisition routines sampled the nasal and oral signals at a rate of 120 Hz. The resultant signal was a ratio of nasal to nasal-plus-oral acoustic energy and was expressed as a Nasalance score.

Figure 5.1 Photograph of the Nasometer model 6200.3, with the headset and computer screen.

5. 2. 3 Pressure/Flow

Nasal flow and pressure and oral pressure were recorded using the PERCI-SARS (Microtronics, Inc., Chapel Hill, NC). The system contained a nasal pressure catheter and an oral pressure catheter which connected to 2 pressure transducers and a nasal flow tube which connected to the pneumotachograph, which was then connected to a flow transducer. The transducers had a response time of 1 ms. The sample rate was set at 1000 and the sample time was set at 10 seconds. This allowed for the recording of nine repetitions of each speech sample on one graph. Oral pressure was measured by placing a catheter in the oral cavity, nasal pressure was measured by securing a nasal catheter with a soft foam cork in one nostril and nasal flow was measured by a heated pneumotachograph connected by plastic tubing in the other nostril (Figure 5.2). Differential pressure was measured by calculating the difference between oral and nasal pressure on either side of the velopharyngeal port. Velopharyngeal port area was calculated using the differential pressure measurements and the rate of nasal flow (Chapter 2. section 5.3).

Figure 5.2 Oral pressure, nasal pressure and nasal flow tubes of the PERCI SARS pressure/flow system in place for data collection.

5. 3 Calibration of Instruments

5. 3. 1 Calibration of the Nasometer

The Nasometer was calibrated after every ten participants in accordance with recommendations of the manufacturers. Calibration was completed by presenting a tone to the microphones on the headset and adjusting the balance between the two microphones (Appendix 14).

5. 3. 2 Calibration of PERCI SARS

The PERCI SARS was calibrated according to the manufacturers instructions at the beginning of each day. The pressure transducers were calibrated by establishing a zero baseline, applying a known pressure of 8 cm H_2O and setting the pressure scale accordingly. Flow was calibrated by establishing a zero baseline, applying a known airflow rate of 250 mls per second and setting the flow scale (Appendix 14).

5. 4 Procedures

All participants were initially assessed for articulation errors using an adapted phonemic screening test (PACS Toys, Grunwell and Harding, 1995, Nfer-Nelson Publishing Co.). Pictures of the toys in this test were used to elicit single word utterances. The participants responses were transcribed phonetically (Appendix 15) and a summary of articulation errors was recorded on the Pacs Toys Phoneme Realisation Chart (Appendix 16).

5. 4. 1 Perceptual assessment of Nasality and Nasal Airflow Errors

This part of the perceptual assessment was made during the live assessment session. The rater sat opposite the child to note any visual characteristics of speech. Ratings of each parameter of nasality and nasal airflow errors were made during the production of the speech sample. Ratings were made at each level of speech - syllable/word; sentence; automatic speech and conversational speech. The following speech sample was obtained (Appendix 12a)[1]:

repetition of syllables 'pa,pa,pa', 'pi,pi,pi' and word 'hamper, hamper, hamper';

16 sentences which the participant repeated after the examiner;

A nursery rhyme (Jack and Jill);

Counting/repetition of 1 to 20 and 60 to 70.

Two minutes of conversational speech.

Nasality and nasal airflow was described subjectively using the classification system developed for this study (see Appendix 13b).

---

[1] High pressure consonants which were vulnerable to nasal airflow errors are underlined in the speech sample.

5. 4. 2 Nasometry assessment

Instrumental measurements of hypernasal and hyponasal resonance were made using the Nasometer. The nasometric assessment was carried out by placing the headset securely on the child's head. The sound separator sat under the nose and above the upper lip without interfering with lip movement (Figure 5.3).

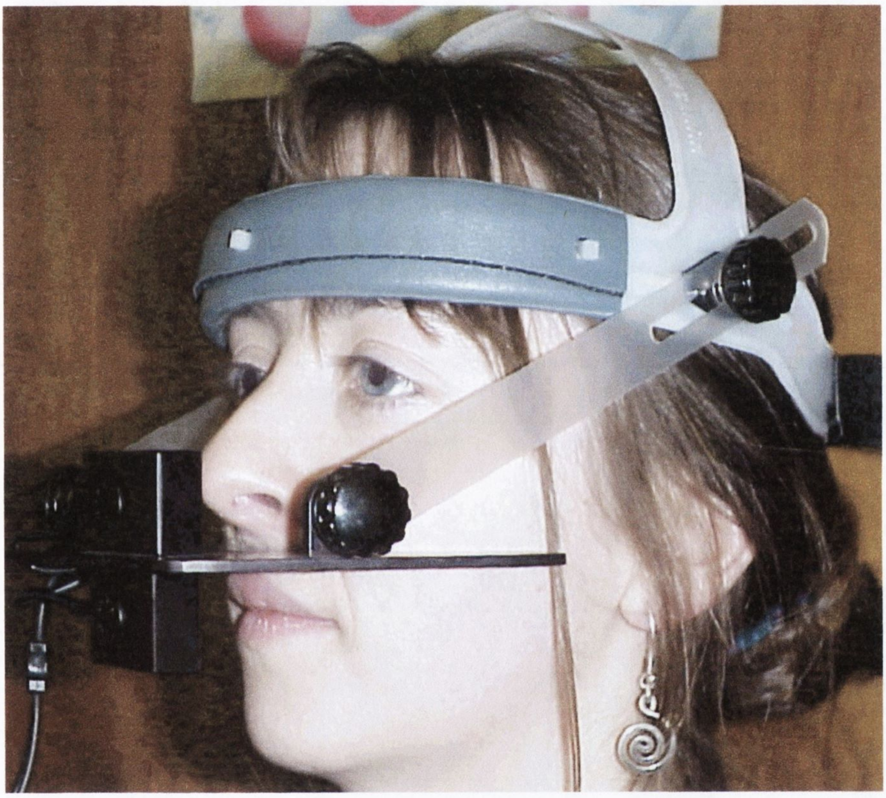


Figure 5.3 Photograph of Nasometer headset in place for data collection

The recording display duration was set at 4 seconds to allow for observation of each sentence category during data collection.. The participant repeated each of the 16 test sentences individually after the examiner. Only the participants production of each sentence was recorded. There was one second between each sentence recording which allowed the examiner to switch on and off the data collection. This was displayed graphically as a space between each sentence. At the end of each category five seconds of silence was recorded in order to identify the separate sentence categories (Figure 5. 4).

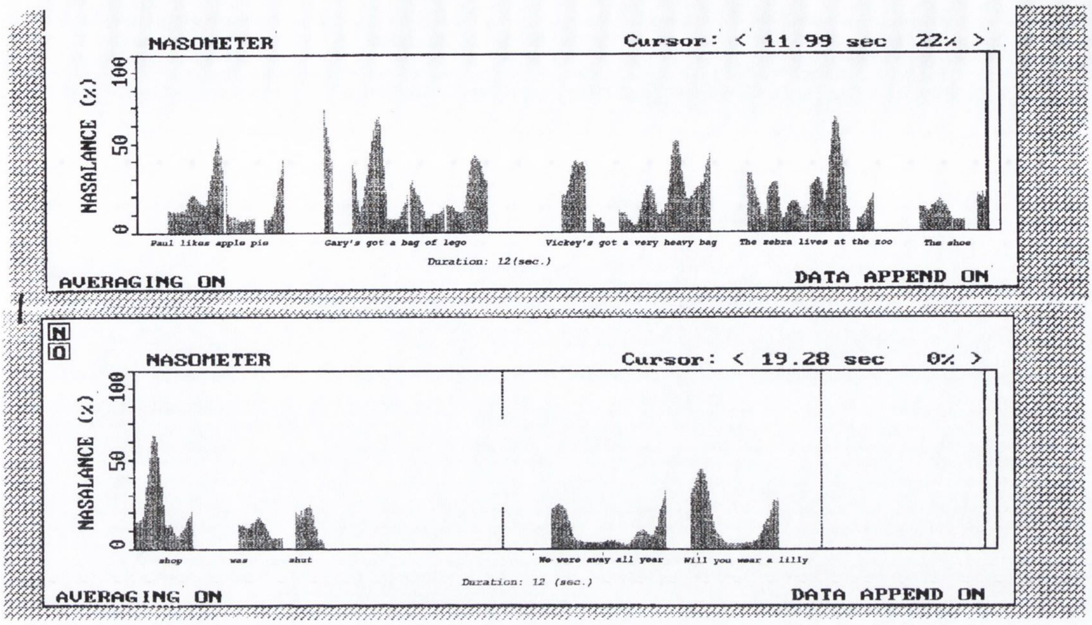


Figure 5.4 Nasometry graphs indicating 1 second gaps between High Pressure consonant sentences in the top graph, and a 5 second gap between the High Pressure consonant sentences and the Low Pressure consonant sentences (between cursors) in the bottom graph.

The sentences were recorded in the following order of categories:

High pressure consonants;

Low pressure consonants;

Mixed consonants;

Nasal consonants (Appendix 7)

The 16 sentences were saved for subsequent analyses.

5. 4. 3 Pressure/Flow assessments

Pressure/flow measurements were made using the PERCI SARS following completion of the nasometric assessment. Before assessing the participant, the system was checked to ensure there was a zero baseline for oral pressure, nasal pressure and nasal flow. Without the tubes being placed in the mouth or nostril, a recording was made and the graphs were saved and reviewed to ensure that there was a recording at zero on the graph. If the baseline was not at zero, calibration was repeated. If the baseline was at zero, the assessment commenced.

Prior to placing the tubes in the child's nostrils, an assessment was carried out to ascertain which nostril was most patent. The assessor pinched the child's right nostril and asked him/her to sniff. This procedure was repeated for the left nostril. The higher the pitch of the sound on inhalation the narrower the nostril. The nostril with the lower pitch sound during inhalation was considered to be the 'clear' nostril (Shprintzen, 1995).

Pressure/flow measurements were made as per instructions in the PERCI SARS Manual (1994). An appropriate sized airflow tube was placed in the child's 'clear' nostril. The pressure catheter was place in the other nostril. The oral pressure catheter was placed in the child's mouth ensuring that the tongue did not obstruct the tube. The child was asked to repeat 'pa pa pa; pa pa pa; pa pa pa' with the tube in place. During sound production the display screen was checked by the examiner to ensure that the baseline was stable. If the baseline was unstable or did not start at 0, the system was recalibrated. The child repeated the speech sample and the data was saved. (Figure 5.5). This procedure was repeated for the speech samples 'pi pi pi' and 'hamper hamper hamper'. Three separate graphs indicating pressure/flow measurements for 'pa', 'pi' and 'hamper' were saved. Simultaneous audio recordings of the speech sample were made to validate analysis.

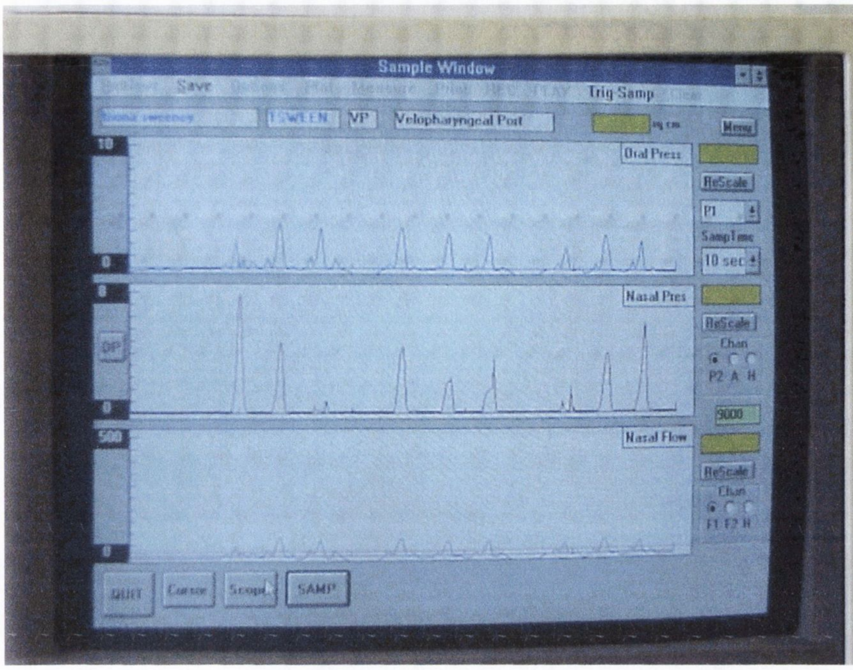


Figure 5.5. Sample window of PERCI SARS pressure/flow analysis during the production of "hamper, hamper, hamper"

5. 5 Instrumental Analyses

Nasometry and pressure/flow analyses were made after the assessment session. Using Nasometry, nasalance data was displayed in 'Contour' mode, with a display duration of 60 seconds. The 16 Total Test sentences were analysed and the nasalance score was computed. The graph of the 16 Total Test sentences was printed. Using a display duration of 20 seconds, the initial 5 sentences containing high pressure consonants were marked with the cursors (Figure 5.6) and a nasalance score was calculated. The Low Pressure consonant sentences, the Mixed consonant sentences, and the Nasal sentence were marked and analysed in the same manner. The mean nasalance score for each speech sample was recorded on the summary form (Appendix 17). (Examples of normal and abnormal graphs are presented in Appendix 18). All silent intervals were included in the analyses as the Nasometery manual states that silence does not affect computed nasalance scores.

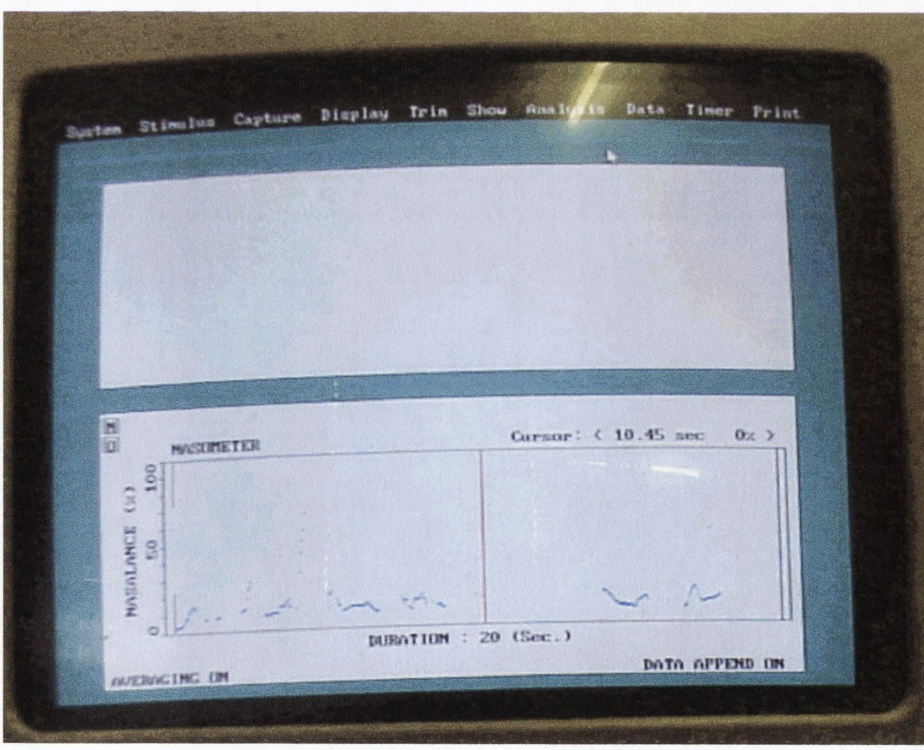


Figure 5.6 Analysis window of Nasometer with cursors marking the beginning and the end of the high pressure consonant sentences.

Pressure/flow measurements were displayed on three channels. The oral pressure was displayed on the topChannel, nasal pressure was displayed on the middle channel and nasal flow was displayed on the bottom channel. Speech playback allowed the audio recording of /p/ to be heard as a cursor moved across the oral pressure graph, permitting accurate identification of the /p/ sound in relation to the oral peak pressure. Pressure/flow measurements were calculated for /p/ in 'pa'. The 'Velopharyngeal Area Measurement' was selected to measure peak oral pressure with associated nasal pressure and nasal flow. The peak oral pressure was selected as this identified the articulatory release of the /p/ sound. (The first production of 'pa' was not used in the analysis as this often has greatest variability in sound production (Van Hattum & Worth, 1967).) Nine repetitions of the syllable /pa/ were recorded, but in order to obtain a mean value for the production of /p/ in 'pa', the next 6 highest oral pressure peaks of this syllable were identified using the cursor (PERCI SARS Manual, 1994). The measurement in the yellow box adjacent to the oral pressure graph confirmed the peak oral pressure (Figure 5.7).

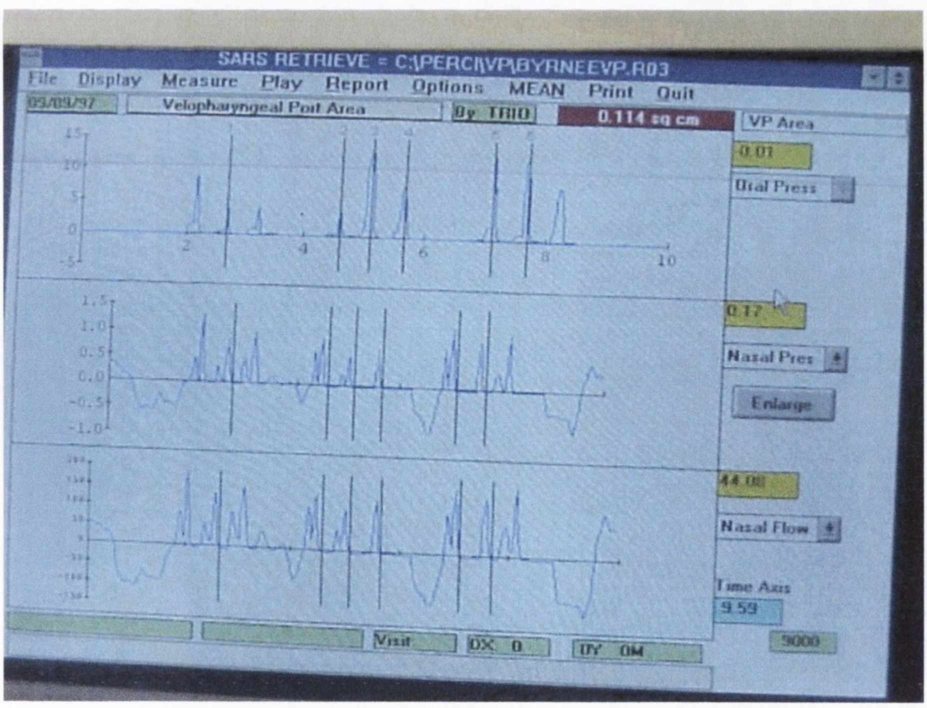
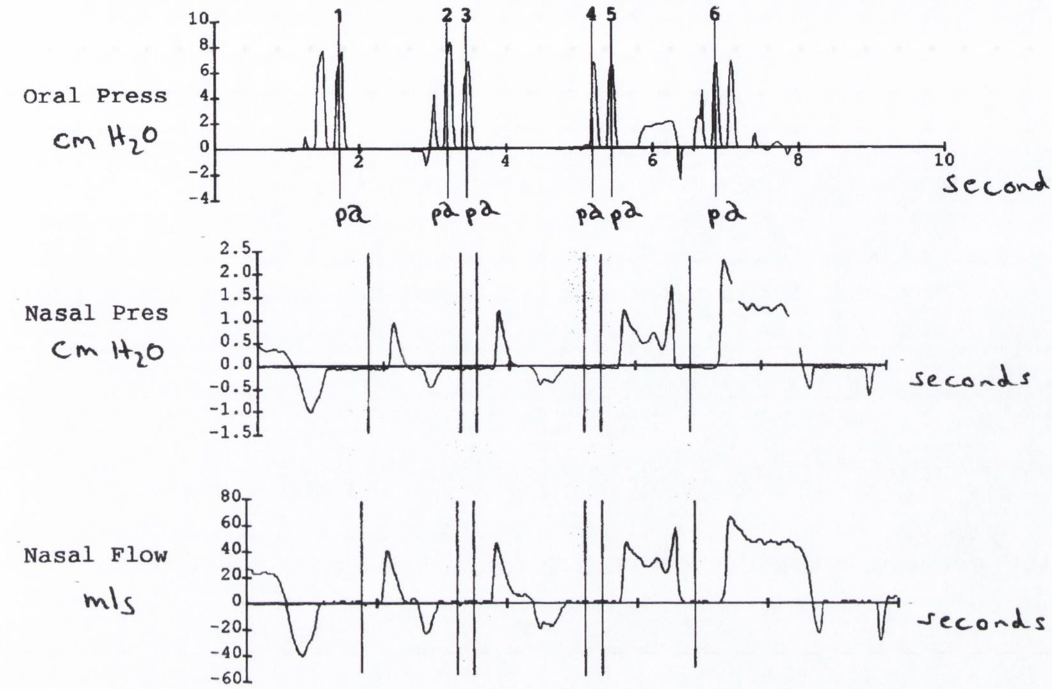


Figure 5.7 Analysis window of PERCI SARS pressure/flow system with cursors placed at peak oral pressure points.

When the oral peaks were selected, simultaneous measurements were made for nasal pressure and nasal flow (Fig 5.8).

'UBLIN CHILDREN'S HOSPITAL SARS REPORT DATE:01-09-96 By:TRIO
Measurement Report From File: K HP.R01 Type: HAMPER WITH VOICE

Patient: Hospital ID: Visit:
DX: 0 Age: 0Y 0M Eval Date: 09/01/96 Examiner: TRIO



Cur#			VP Area				
1	OP=07.6	NP=-00.1	NFlow=000.0	SPL=061.0	DP=07.63	AREA= 0.000	X=1.7
2	OP=08.3	NP=-00.1	NFlow=001.8	SPL=060.0	DP=08.37	AREA= 0.001	X=3.2
3	OP=06.8	NP=-00.1	NFlow=000.0	SPL=060.0	DP=06.80	AREA= 0.000	X=3.4
4	OP=06.7	NP=-00.1	NFlow=000.0	SPL=061.0	DP=06.77	AREA= 0.000	X=5.2
5	OP=06.5	NP=-00.1	NFlow=001.8	SPL=061.0	DP=06.50	AREA= 0.001	X=5.4
6	OP=06.6	NP=-00.1	NFlow=000.0	SPL=061.0	DP=06.63	AREA= 0.000	X=6.8

Oral Press	Mean = 7.1	STD = 00.73	SE= 00.30
Nasal Pres	Mean = -0.1	STD = 00.01	SE= 00.00
Nasal Flow	Mean = 0.6	STD = 00.92	SE= 00.37
SPL	Mean = 60.7	STD = 00.52	SE= 00.21
Computed DP	Mean = 7.12	STD = 00.73	SE= 00.30
AREA(sq cm)	Mean = 0.000	STD = 0.000	SE= 0.000

Figure 5.8. PERCI SARS graph with cursors number (Cur#) 1 to 6, indicating peak oral pressure and corresponding nasal pressure and nasal flow for a normal speaker, during the production of the syllable 'pa'. (There is a time shift of 0.5 seconds for the nasal pressure and nasal flow graphs due to the graphic software used). Measurements for oral pressure (OP), nasal pressure (NP), nasal flow (Nflow), sound pressure level (SLP), differential pressure (DP) and velopharyngeal port area (AREA) are presented for each cursor. The mean measurements, with standard deviation and standard error are presented in the bottom section of the graph.

Differential pressure and velopharyngeal port area were automatically calculated. Subsequently mean measurements were computed for all 6 oral pressure, nasal pressure, nasal flow, differential pressure and velopharyngeal area measurements (Examples of normal and abnormal PERCI SARS graphs are presented in Appendix 19). The analysis was then repeated for the /p/ in 'pi' and subsequently for /p/ in 'hamper'.

On completion of analyses, a Summary Report Form was completed. This included: the perceptual ratings; nasalance scores for Total Test sentences, High Pressure consonant sentences, Low Pressure consonant sentences, mixed sentences and the nasal sentences; oral pressure measurements, differential pressure measurements, nasal flow measurements and velopharyngeal port area measurements for /p/ in 'pa', 'pi' and 'hamper' and for /m/ in 'hamper'; an articulation summary, voice quality, and information on orofacial features, hearing, language and cognitive development (Appendix 17).

5. 6 Assignment of Scales

Analysis of each of the perceptual parameters was made under the following headings:

hypernasality
hyponasality
nasal emission
nasal fricative
nasal turbulence
velopharyngeal fricative.

Descriptors were assigned a numerical value to allow for statistical analysis.
Hypernasality was rated as 0 if absent and a numerical value as follows :

1 = a) mild, evident but acceptable;
2 = b) mild/moderate, unacceptable distortion evident on high vowels;
3 = c) moderate, evident on high and low vowels;
4 = d) moderate/severe, evident on all vowels and some consonants;
5 = e) severe, evident on all vowels and most voiced consonants.

Hyponasality was rated 0 if absent and as follows:

1 = a) evident, but acceptable;
2 = b) moderate, reduced nasality on all vowels and some consonants;

3 = c) severe, total denasal production of nasal consonants.

Consistency was recorded as 1 = inconsistent and 2 = consistent.
Lower numbers indicated a milder degree of abnormality.

Each category of nasal airflow was grouped descriptively according to strength, consistency, frequency:

weak, inconsistent, infrequent	=	1
weak, inconsistent, frequent	=	2
weak, consistent, infrequent	=	3
weak, consistent, frequent	=	4
strong, inconsistent, infrequent	=	5
strong, inconsistent, frequent	=	6
strong, consistent, infrequent	=	7
strong, consistent, frequent	=	8

The numerical values ranged from 1 (indicating a mild problem) to 8 (indicating a severe problem).

5.7 Statistical Analysis

The distribution of the perceptual results was calculated using SPSS statistical analysis programme. The number of cases and the percentage of the study group who presented with varying degrees of abnormality were found.

Reliability studies were carried out for all test procedures. Reliability of the Perceptual Profile has been presented in Chapter 4. Reliability of the instrumentation was calculated using generalizability theory, which states that in any measurement situation there are multiple sources of error variance. To obtain the generalizability reliability coefficient the mean square for the 10 participants, the speech sample and the time were calculated using a repeated measures analysis of variance. The estimated mean square was calculated. The reliability coefficient was calculated using the following equation:

$$\frac{\rho^2_{p} + \rho^2_{ps}}{\rho^2_{p} + \rho^2_{pt} + \rho^2_{ps} + \rho^2_{pst}}$$

Where ρ^2 is the estimate mean square, p is the participant, s is the speech and t is the time.

The Nasometer provided statistical data on mean nasalance scores for the speech samples. PERCI SARS provided mean nasal airflow (mls/sec), mean oral pressure (cm H_2O), mean nasal pressure (cm H_2O), mean differential pressure (cm H_2O) and mean velopharyngeal port area (cm^2) for /p/ in the three speech samples.

Mean, range and standard deviation of instrumental data from the study group was calculated using the SPSS statistical analysis programme. Differences between normal and abnormal speakers were assessed using raw data from the normal nasalance study for Irish children (Chapter 3). Repeated measures analysis of variance ($\alpha = .05$) were carried out to analyse the effect of the following three interactions: normal speakers, pathological speakers and sentence category. A Bonferroni post-hoc analysis was carried out ($\alpha = .05$) to examine the differences between normal and abnormal speakers during the production of each speech category.

Raw data of pressure/flow measurements for normal speakers was not available; however, the mean and standard deviation of a normal group was known (Zajac, 1998, personal communication)[2]. Due to the large sample size it was possible to carry out independent samples t tests to ascertain whether or not the study group results on the pressure/flow system differed significantly from results of normal speakers. The tests were two tailed ($\alpha = .05$). The following formula was used to obtain t values:

[2] Zajac et al. (1997) provided normative data for three groups of children categorised according to age. For the purpose of this study, mean and standard deviation scores of the entire group were provided by Zajac (1998, personal communication).

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\text{standard error of difference}}$$

Where standard error of differences = $\sqrt{\dfrac{S_1}{N_1} + \dfrac{S_2}{N_2}}$ and df = $n_1 + n_2 - 2$

Ratings of nasality were considered as interval scales. Hence, the Pearson product moment correlation was performed using the SPSS package to examine the relationship between nasalance scores and perceptual ratings of nasality. Ratings of nasal airflow were considered ordinal in nature. Hence, the Spearman rank correlation was performed using the SPSS package to ascertain the relationship between perceptual ratings of nasal airflow and pressure/flow measurements. All correlational analyses were two tailed (α =. 05).

Test sensitivity, specificity and overall efficiency were calculated by comparing instrumental results with perceptual data. Test sensitivity is the percentage of participants that are identified as having abnormal speech on one test and who are also identified as having abnormal speech on another test. Test specificity refers to the percentage of patients who are identified as having normal speech on one test and who are also identified as having normal speech on another test. An overall efficiency rating can be calculated adding the number of times one test agreed with the other test, divided by the total number of opportunities (i.e. total number of participants tested) (Watterson et al., 1998).

CHAPTER 6

RESULTS

6.1 Participants

A consecutive series of 50 participants attending a Cleft Palate Unit for speech problems associated with cleft palate, velopharyngeal dysfunction and nasal obstruction were assessed. There were 30 males and 20 females with a mean age of 9;5 years. Ages ranged from 4;10 years to 15;10 years. The diagnosis of each participant was recorded either at the time of assessment (as in the cases of obvious cleft palate and nasal obstruction), following the assessment session (which included perceptual, instrumental and oral assessment with a detailed case history) or following further team assessments (Table 6.1).

Table 6.1 Diagnosis for the 50 participants in the study group.

Classification	Number of cases	Percent
Velopharyngeal dysfunction	24	48%
Secondary cleft palate	13	26%
Submucous cleft palate	4	8%
Unilateral cleft lip & palate	3	6%
Bilateral cleft lip & palate	2	4%
Nasal obstruction	2	4%
Velocardiofacial syndrome	1	2%
Adenoidal hypertrophy	1	2%

Of the twenty four participants presenting with velopharyngeal dysfunction, two had velopharyngeal mislearning and one had velopharyngeal incompetency associated with a mild dyspraxia.

6. 2 Perceptual Results

6. 2. 1 Reliability of Perceptual Assessments

Reliability of the Perceptual Profile was assessed prior to commencement of data collection. Details of the reliability studies are presented in Chapter 4 Sections 4.4 - 4.8. Results indicated good to excellent intra-rater reliability for the rater in this main study. However inter-rater reliability varied between raters and in sections of the Profile: nasal fricatives; velopharyngeal fricatives, mild nasal emission and mild nasal turbulence. The percent of agreement between raters was acceptable in all studies.

6. 2. 2 Nasality

Perceptual ratings of nasality were divided into two categories, hypernasality and hyponasality, using two separate descriptive scales. Nasality was rated as consistent, if the degree of nasality was the same for each level of speech (words, sentences, automatic speech, conversational speech). Nasality was rated as inconsistent if the degree of nasality varied across speech levels.

Hypernasality

Perceptual ratings of hypernasality ranged from absent (0) to severe (e), with a mean rating of mild/moderate (unacceptable distortion evident on high vowels). The perceptual ratings of hypernasality are summarised below (Table 6.2). Of the 40 participants (80%) who were perceived as hypernasal, 27 (67%) had inconsistent hypernasality and 13 (33%) had consistent hypernasality.

Table 6.2 Perceptual ratings of hypernasality for the study group, with grading and description of hypernasality.

Rating	**Description**	**Number of cases**	**Percent**
0	absent	10	20%
a - mild	evident but acceptable	4	8%
b - mild/moderate	unacceptable distortion, evident on high vowels	19	38%
c - moderate	evident on high & low vowels	8	16%
d - moderate/severe	evident on all vowels & some consonants	6	12%
e - severe	evident on all vowels & most voiced consonants	3	6%

Hyponasality

Seven participants (14%) were perceived as hyponasal, two had mild hyponasality (evident but acceptable), and five had moderate hyponasality (reduced nasality on all vowels and some nasal consonants) (Table 6.3). Hyponasality was inconsistent for two participants.

Table 6.3 Perceptual ratings of hyponasality for the study group. Descriptors of ratings, number of cases and percentage of the group is presented.

Rating	**Description**	**Number of cases**	**Percent**
0	absent	43	86%
a - mild	evident but acceptable	2	4%
b - moderate	evident on all vowels and some consonants	5	10%

6. 2. 3 Nasal Airflow

For the purpose of statistical analysis, the ratings for each category of nasal airflow (nasal emission, nasal fricative, nasal turbulence, velopharyngeal fricative) were ranked from 1 to 8 according to strength, consistency and frequency (Chapter 5, Section 6).

Nasal Emission

Nasal emission was perceived in 27 participants (54%) and ranged from weak, inconsistent and infrequent to strong, consistent and frequent (Table 6.4).

Table 6.4 Perceptual ratings of nasal emission. Number and percent of participants presenting with varying degrees of nasal emission.

Description	Number of cases	Percent
Absent	23	46%
weak, inconsistent, infrequent	11	22%
weak, consistent, infrequent	3	6%
weak, consistent, frequent	8	16%
strong, consistent, frequent	5	10%

Nasal fricative

Eight participants (16%) were perceived as having nasal fricatives during assessment (Table 6.5). Seven out of eight participants had phoneme specific nasal fricatives.

Table 6.5 Perceptual ratings of nasal fricatives. Number and percent of participants with varying degrees of nasal fricatives.

Description	Number of cases	Percent
Absent	42	84%
weak, inconsistent, infrequent	3	6%
weak, consistent, frequent	1	2%
strong,consistent, infrequent	2	4%
strong, consistent, frequent	2	4%

Nasal turbulence

Seventeen participants (34%) had nasal turbulence, ranging from weak, inconsistent and infrequent to weak, consistent and frequent (Table 6.6). None of the participants had strong nasal turbulence.

Table 6.6 Perceptual ratings of nasal turbulence. Number and percent of participants presenting with varying degrees of nasal turbulence.

Description	Number of cases	Percent
Absent	33	66%
weak, inconsistent, infrequent	7	14%
weak, inconsistent, frequent	1	2%
weak, consistent, infrequent	3	6%
weak, consistent, frequent	6	12%

Velopharyngeal fricative

Seven participants (14%) had velopharyngeal fricatives and six of these were phoneme specific. Five participants had strong, consistent and frequent velopharyngeal fricatives (Table 6.7).

Table 6.7 Perceptual ratings of velopharyngeal fricative. Number and percent of participants with varying degrees of velopharyngeal fricative.

Description	Number of cases	Percent
Absent	43	86%
weak, inconsistent, infrequent	2	4%
strong, consistent, frequent	5	10%

6. 3 Instrumental Results

6. 3. 1 Nasometry - Reliability

Test-retest reliability of the nasometer was assessed during the study on nasalance scores for normal English-speaking Irish children. Details are reported in Chapter 3, Section 2. Results indicated good test-retest reliability for the Total Test sentences, High Pressure consonant sentences and Low Pressure consonant sentences. However, one participant had a 10% difference in nasalance scores for the Nasal sentence across the two assessments. If this participant was removed from the reliability calculations, the reliability coefficient for the Nasal sentence improved from 0.15 to 0.79.

6. 3 2 Nasometry Results

Mean nasalance scores, standard deviations and range of scores were calculated for the study group during the production of the Total Test sentences, sentences containing High Pressure Consonants (devoid of low pressure consonants and nasals), sentences containing Low Pressure consonants (devoid of high pressure consonants and nasals), Mixed Consonants sentences (containing high pressure consonants, low pressure consonants and nasals) and a Nasal sentence (containing 50% nasal consonants) (Table 6.8). (Examples of graphs are presented in Appendix 18).

Table 6.8 Mean nasalance scores, standard deviation and range of scores for each speech sample. Scores for normal speakers of similar age and dialect are presented in brackets.[1]

Speech sample	**Mean** ***(normal)***	**SD** ***(normal)***	**Range** ***(normal)***
Total Test sentences	40% *(25%)*	15% *(5%)*	5% - 68% *(15% - 35%)*
High pressure consonant sentences	33% *(14%)*	17% *(5%)*	4% - 67% *(7% - 25%)*
Low pressure consonant sentences	38% *(16%)*	18% *(6%)*	4% - 64% *(7% - 30%)*
Mixed consonant sentences	43% *(34%)*	15% *(7%)*	5% - 70% (19% - 47%)
Nasal sentence	53% *(51%)*	17% *(7%)*	6% - 79% *(33% - 68%)*

Analysis of the median scores, the number of participants in the 25th and 75th percentiles, and range of scores for each group of speakers indicates substantial differences between normal and abnormal speakers (Figure 6.1).

[1] The normal scores for the Total Test sentences included the 70 participants from the normal Nasometry study (Chapter 3) and 26 participants from the Nasometry pilot study 2 (Appendix 5). The larger number of participants produced a slightly lower mean nasalance score of 25% (26.6% in smaller group), similar standard deviation scores and a slightly wider range 15% - 35% (17% - 35% in smaller group).

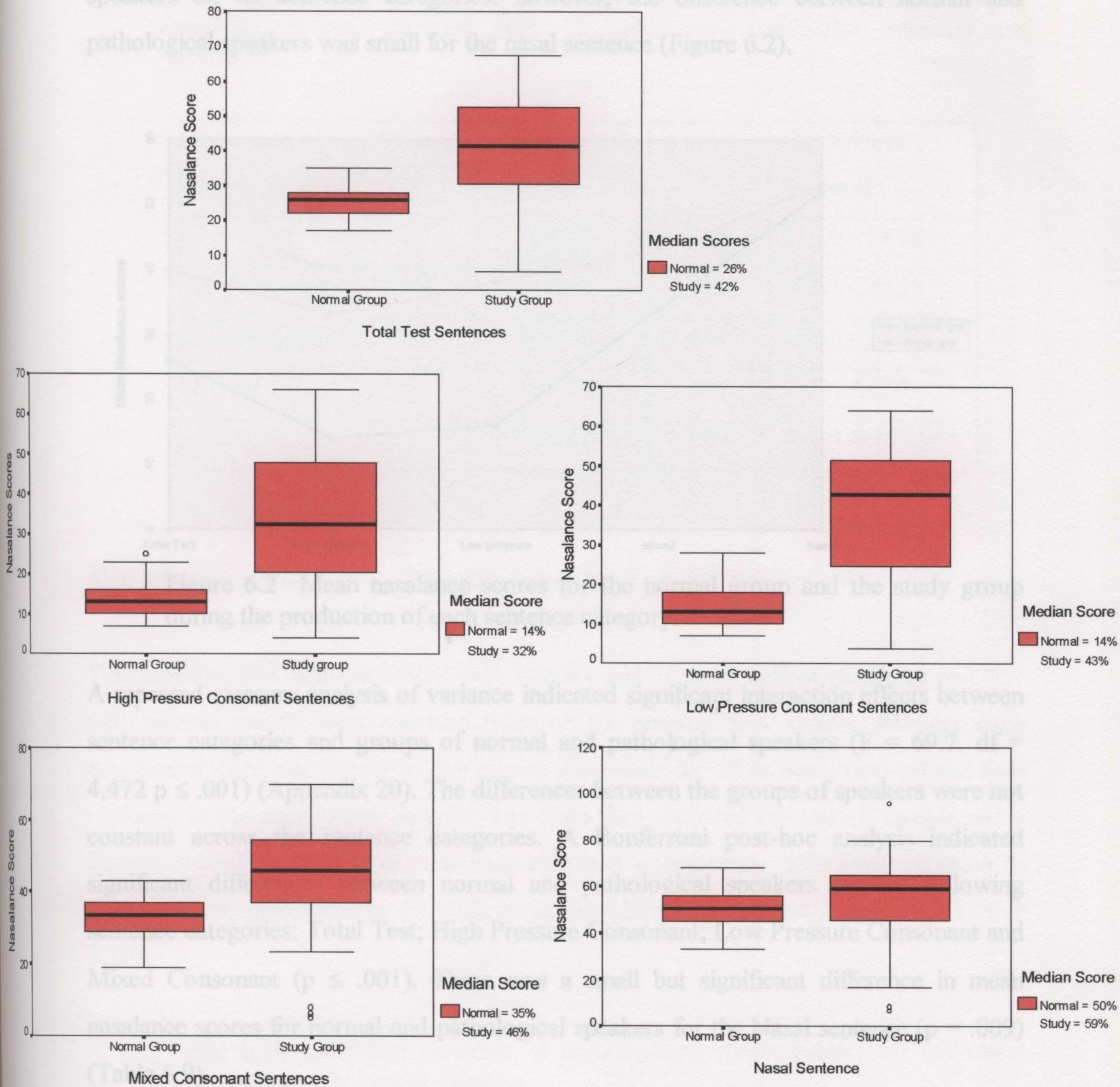


Figure 6.1 Mean nasalance scores for each sentence category (thick black line) and the 50% of participants who fall between the 25th and 75th percentile (red box). The highest and lowest values (excluding outliers and extremes) is shown by thin black lines. The outliers (o) indicate values between 1.5 and 3 box-lengths from the upper edge of the box.

There were significant differences between mean nasalance scores for each group of speakers for all sentence categories; however, the difference between normal and pathological speakers was small for the nasal sentence (Figure 6.2).

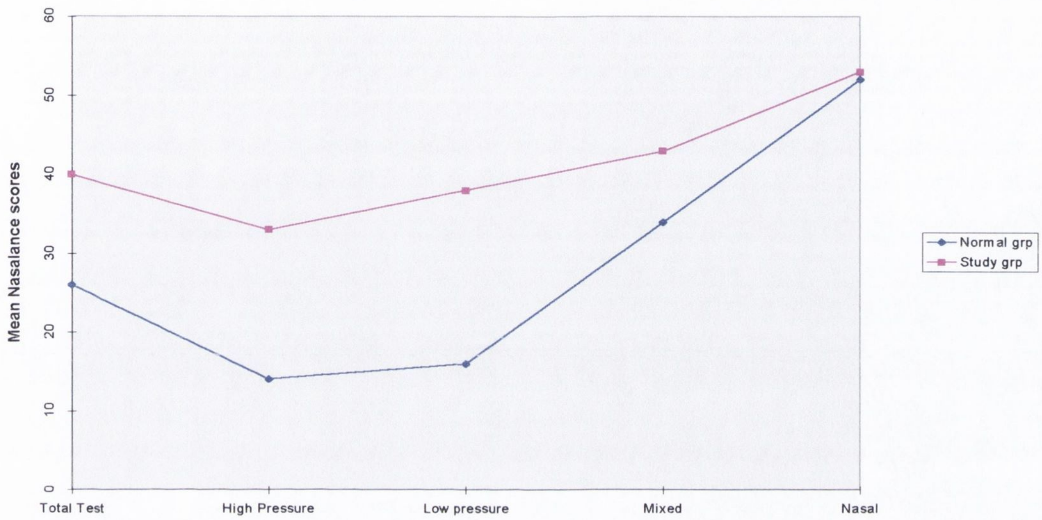


Figure 6.2 Mean nasalance scores for the normal group and the study group during the production of each sentence category.

A repeated measure analysis of variance indicated significant interaction effects between sentence categories and groups of normal and pathological speakers ($F = 69.7$, $df = 4,472$ $p \leq .001$) (Appendix 20). The differences between the groups of speakers were not constant across the sentence categories. A Bonferroni post-hoc analysis indicated significant differences between normal and pathological speakers for the following sentence categories: Total Test; High Pressure Consonant; Low Pressure Consonant and Mixed Consonant ($p \leq .001$). There was a small but significant difference in mean nasalance scores for normal and pathological speakers for the Nasal sentence ($p = .009$) (Table 6.9).

Table 6.9 Bonferroni post-hoc analysis of the difference between normal and pathological speakers for each sentence category. Differences in mean nasalance scores, standard error, p value and confidence intervals are presented.

Sentence Type	**Difference**	**Standard error**	**p value**	**Confidence Interval**
Total Test	13.52	0.9158	$\leq .001$	11.16 - 15.88
High Pressure	18.19	0.9158	$\leq .001$	15.83 - 20.55
Low Pressure	21.90	0.9158	$\leq .001$	19.54 - 24.26
Mixed	9.17	0.9158	$\leq .001$	6.81 - 11.53
Nasal	2.37	0.9158	$= .009$	0.01 - 4.73

6. 3. 3 Pressure/Flow (PERCI SARS) Reliability

Ten children, five who presented with nasality and/or nasal airflow problems and five with normal speech, were assessed twice on the pressure/flow system. Between each assessment there was a three to five minute gap, while the tubes were removed and then replaced in the appropriate nostrils. The child then repeated the speech sample and a second recording was taken. Analysis indicated considerable individual variation in pressure/flow measurements for pathological speakers between test 1 and test 2 (Table 6.10).

Table 6.10 Individual variation in pressure/flow measurements on test-retest analyses for pathological speakers.

Pressure/Flow Measurement	**Range of difference between Test 1 & Test 2**
Oral Pressure	0.2 cm H_2O to 2.9 cm H_2O
Nasal Flow	0.5 mls to 40 mls
Differential Pressure	0.2 cm H_2O to 3.4 cm H_2O
Velopharyngeal Port Area	0 cm^2 to .3 cm^2

Results of test-retest analyses for the five normal speakers indicated similar variation in oral and differential pressure measurements during the two assessments, but smaller variations in nasal flow and velopharyngeal port area measurements (Table 6.11).

Table 6.11 Individual variation in pressure/flow measurements on test-retest analyses for normal speakers.

Pressure/Flow Measurement	**Range of difference between Test 1 & Test 2**
Oral Pressure	0.2 cm H_2O to 2.5 cm H_2O
Nasal Flow	0.5 mls to 9 mls
Differential Pressure	0.2 cm H_2O to 3.4 cm H_2O
Velopharyngeal Port Area	0 cm^2 to 0.007 cm^2

Reliability coefficients were calculated for test-retest scores for the four pressure/flow measurements, using Generalizability Theory (Streiner & Norman, 1995) (Table 6.12). Analysis indicated excellent reliability for differential pressure measurements, good reliability for nasal flow and velopharyngeal port area (VPA) and weak reliability for oral pressure measurements. Scattergrams of test-retest results for the pressure/flow measurements are presented in Appendix 21.

Table 6.12 Test-retest reliability for pressure/flow measurements.

	Oral Pressure	**Nasal Flow**	**Differential Pressure**	**VPA**
Reliability Coefficient	0.66	0.88	0.90	0.77

6. 3. 4 Pressure/Flow Results

Pressure/flow measurements for the production of /p/ in 'pa', 'pi' and 'hamper' included nasal airflow, oral pressure, differential pressure and velopharyngeal port area (Example graphs are presented in Appendix 19). Results of the measurements for each speech sample are presented. One participant had unstable baselines for pressure and flow measurements on all three speech stimuli, another had unstable baseline for /p/ in 'pi' and 'hamper' while a third participant had an unstable baseline for /p/ in 'pi'. This rendered these measurements unreliable and they were excluded.

6. 3. 4 (i) Nasal Flow Measurements

Nasal flow was measured during the production of /p/ in 'pa', 'pi' and 'hamper'. The mean nasal flow (ml/s), the range and standard deviation of nasal flow measurements for each production of /p/ were calculated (Table 6.13).

Table 6.13 Mean nasal flow, standard deviation and range of nasal flow measurements in millilitres per second for the study group during the production of /p/ in 'pa', 'pi' and 'hamper'. Nasal flow measurements for a group of normal speakers of similar age are presented in italic brackets (Zajac, 1998, personal communication).

Speech sample (number)	**Mean ml/s *(normal)***	**SD ml/s *(normal)***	**Range ml/s *(normal)***
pa (49)	48.8 *(1.2)*	77 *(1.41)*	0 - 406 *(0 - 9)*
pi (47)	47.5 *(1.26)*	70 *(1.58)*	0 - 307 *(0 - 13)*
mp (48)	66.3 *(4)*	79 *(5.6)*	0 - 315 *(-1 - 39)*

Results indicated a wide range of nasal flow measurements in the study group (Figure 6.3). The outliers indicate severely abnormal nasal flow during production of /p/.

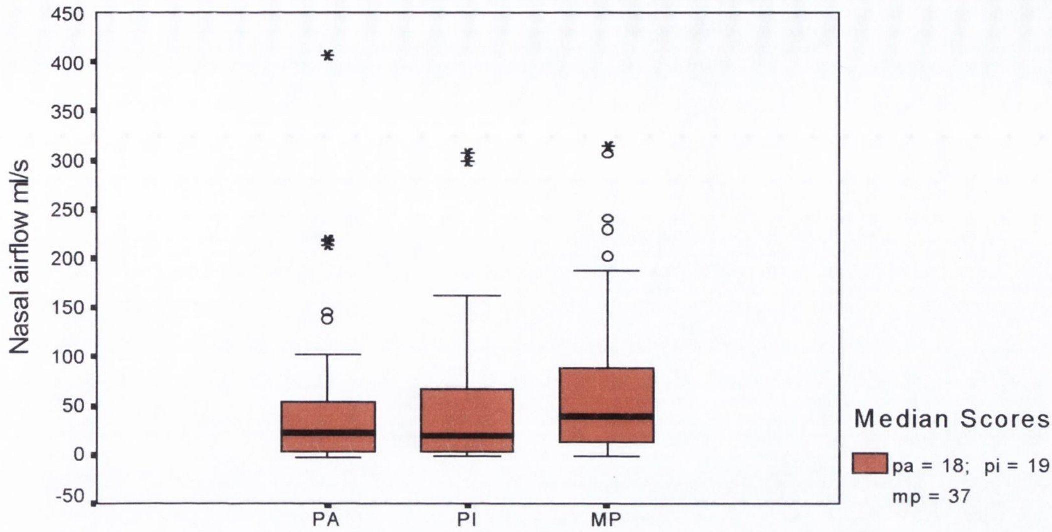


Figure 6.3. The distribution of nasal airflow during the production of /p/ in the 3 different phonetic contexts. The graph indicates the median nasal flow measure (thick black line) and the 50% of participants who fall between the 25th and 75th percentile (red box). The highest and lowest values (excluding outliers and extremes) are shown by thin black lines. The outliers (o) indicate values between 1.5 and 3 box-lengths from the upper edge of the box and the extremes (*) indicate values that are greater than 3 box-lengths from the upper edge of the box.

Independent sample t-tests were used to compare nasal flow results from the study group with nasal flow results from normal speakers of similar age obtained by Zajac (1998, personal communication). There were significant differences between nasal flow measurements for the study group and normal speakers during /p/ in 'pa' ($t = 7.65$, $df = 199$, $p \leq .001$, 95% CI = 26,69), 'pi' ($t = 8.19$, $df = 197$, $p \leq .001$, 95% CI = 26,66) and 'hamper' ($t = 9.72$, $df = 198$, $p \leq .001$, 95% CI = 52,72) (Figure 6.4).

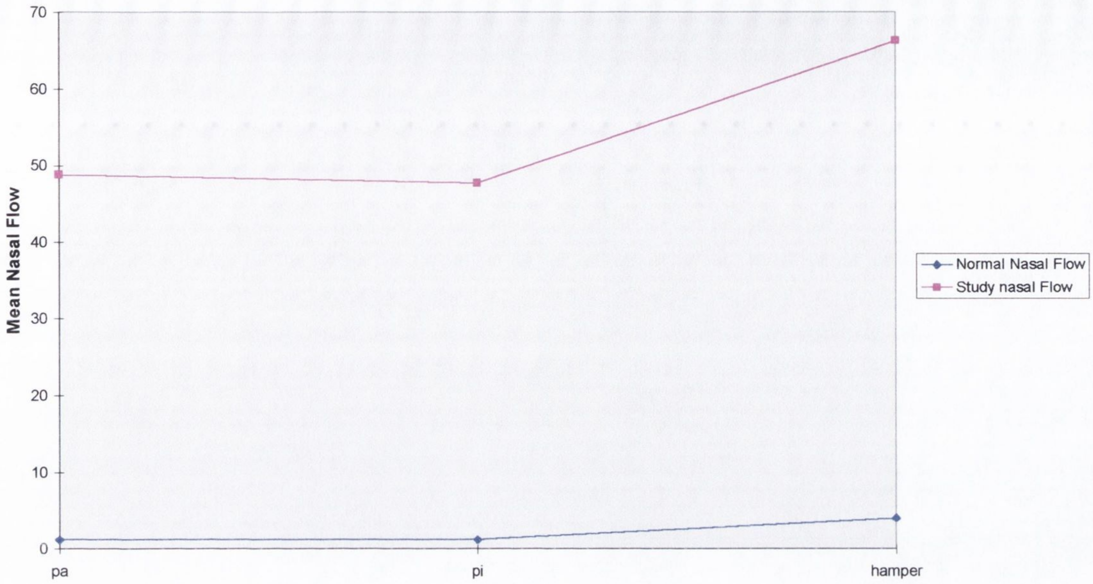


Figure 6.4 Mean nasal flow measurements for normal speakers and the study group during the production of the three speech stimuli.

6. 3. 4 (ii) Oral Pressure Measurements

Oral pressure measurements for /p/ during the production of 'pa', 'pi' and 'hamper' were calculated. Mean oral pressure measurements, standard deviations and range of measurements for the group were found (Table 6.14).

Table 6.14 Mean oral pressure, standard deviation and range in CM H_2O during the production of /p/ in 'pa', 'pi' and 'hamper'. Measurements for a group of normal speakers of similar age are presented in italic brackets (Zajac, 1998, personal communication).

Speech Sample (number)	**Mean** CM ***(normal)***	**SD** CM H_2O ***(normal)***	**Range** CM H_2O ***(normal)***
pa (49)	6.1 *(7.8)*	2.7 *(2)*	1.8 -14.4 *(4.1 - 15.3)*
pi (47)	6.8 *(7.7)*	2.7 *(2.1)*	.9 - 13.8 *(4.1 - 14.9)*
mp (48)	5 *(7)*	2.3 *(2.1)*	.6 - 11.2 *(2.6 - 16.3)*

Results indicate a wide range of oral pressure measurements, with the mean and median value above 3 cm H_2O (i.e. within normal limits) (Figure 6.5).

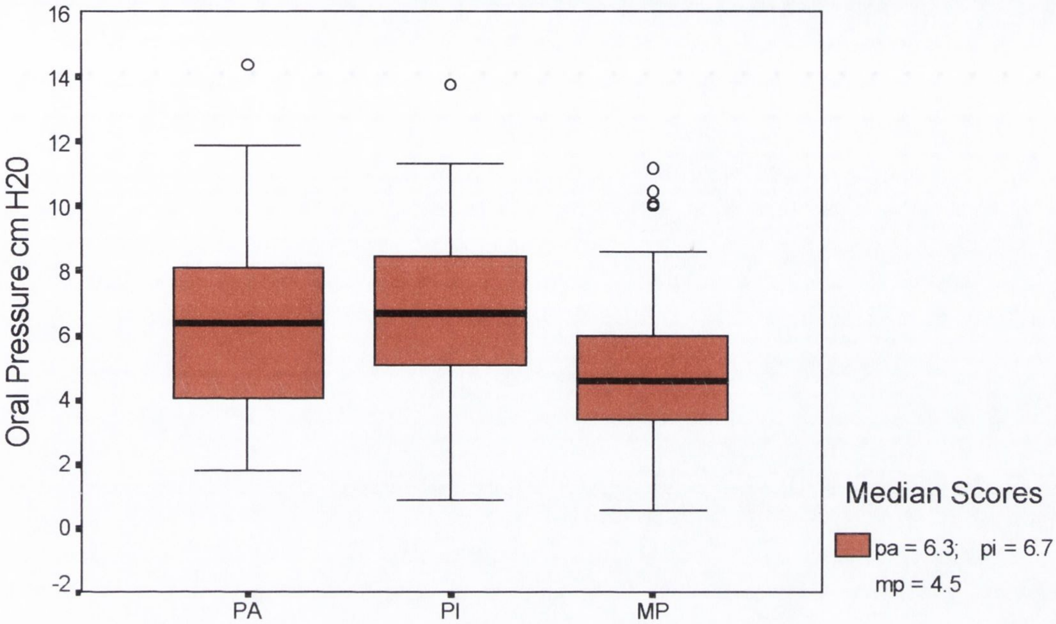


Figure 6.5 The distribution of oral pressure measurements for each speech sample. The graph indicates the median nasal flow measure (thick black line) and the 50% of participants who fall between the 25th and 75th percentile (red box). The highest and lowest values excluding outliers and extremes are shown by thin black lines. The outliers (o) indicate values between 1.5 and 3 box-lengths from the upper edge of the box.

Independent sample t-tests were used to compare oral pressure measurements for the study group with oral pressure measurements for normal speakers (Zajac, 1998, personal communication). Results indicated small but significant differences between measurements for the two groups for the speech stimuli 'pa' ($t = 4.7$, $df = 199$, $p < .001$, 95% CI = 0.92,2.4), 'pi' ($t = 2.5$, $df = 198$, $p = .02$, 95% CI = 0.12,1.68) and 'hamper' ($t = 5.6$, $df = 198$, $p < .001$, 95% CI = 1.28,2.72). (Figure 6.6).

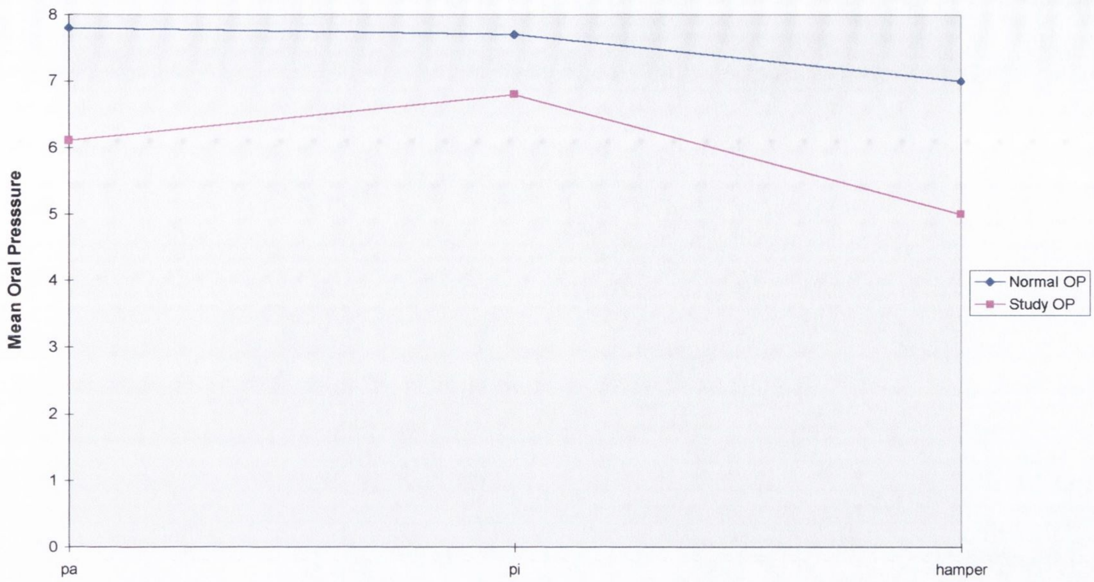


Figure 6.6 Mean oral pressure measurements for normal speakers and the study group during the production of the three speech stimuli.

6. 3. 4 (iii) Differential Pressure Measurements

Mean differential pressure, standard deviation and range of measurements for the study group were calculated using the PERCI SARS software (Table 6.15). No normal values are available and therefore no comparisons can be made. However, normal differential pressures are reported to be greater than or equal to 3 cm H_2O (Warren, 1979).

Table 6.15 Mean differential pressure, standard deviation and range of measurements in cm H_2O during the production of /p/ in 'pa', 'pi' and 'hamper'.

Sample (n)	**Mean** CM H_2O	**SD** CM H_2O	**Range** CM H_2O
pa (49)	5.6	3.4	.03 - 14.5
pi (47)	6.4	3.8	.03 - 19.2
mp (48)	3.2	3.2	.03 - 11.2

Analysis of the 25th to 75th percentile and the range of values indicate that differential pressure measurements were above 3 cm H_2O for a large proportion of the study group (Figure 6.7).

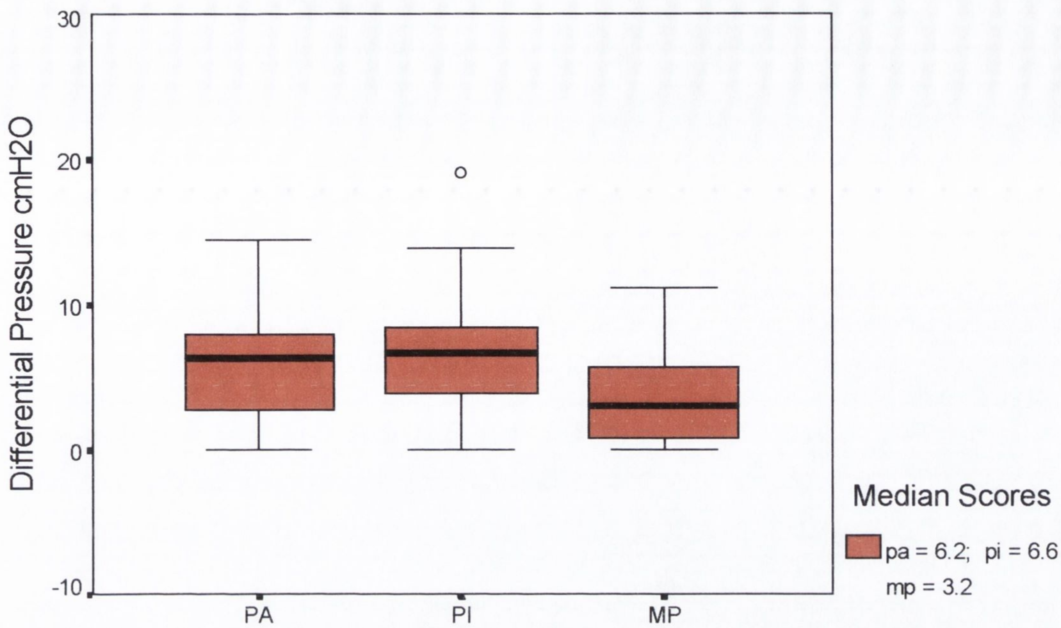


Figure 6.7 Distribution of differential pressure scores during the production of /p/ in 'pa', 'pi' and 'hamper'. The graph indicates the median nasal pressure measure (thick black line) and the 50% of participants who fall between the 25th and 75th percentile (red box). The highest and lowest values (excluding outliers and extremes) are shown by thin black lines. The outliers (o) indicate values between 1.5 and 3 box-lengths from the upper edge of the box.

6. 3. 4 (iv) Velopharyngeal Port Area

Velopharyngeal port area was calculated by the PERCI SARS software package, using the differential pressure between oral pressure and nasal pressure and nasal airflow. Mean velopharyngeal port area, standard deviations and range of measurements for the study group were calculated for each speech stimulus (Table 6.16).

Table 6.16 Mean velopharyngeal port area, standard deviations and range of area measurements for the study group during the production of /p/ in the 3 samples. Measurements for a group of normal speakers of similar age are presented in italic brackets (Zajac, 1998, personal communication).

Speech Sample (number)	**Mean cm^2 *(normal)***	**SD cm^2 *(normal)***	**Range cm^2 *(normal)***
pa (49)	.083 *(.0005)*	.189 *(.0006)*	0 - 0.8 *(0 - .004)*
pi (47)	.077 *(.0005)*	.156 *(.0008)*	0 - 0.69 *(0 - .006)*
mp (48)	.137 *(.0019)*	.219 *(.0026)*	0 - 0.79 *(0 - .022)*

Results indicated a wide range of velopharyngeal area measurements, with many participants having large velopharyngeal port area measurements (Figure 6.8).

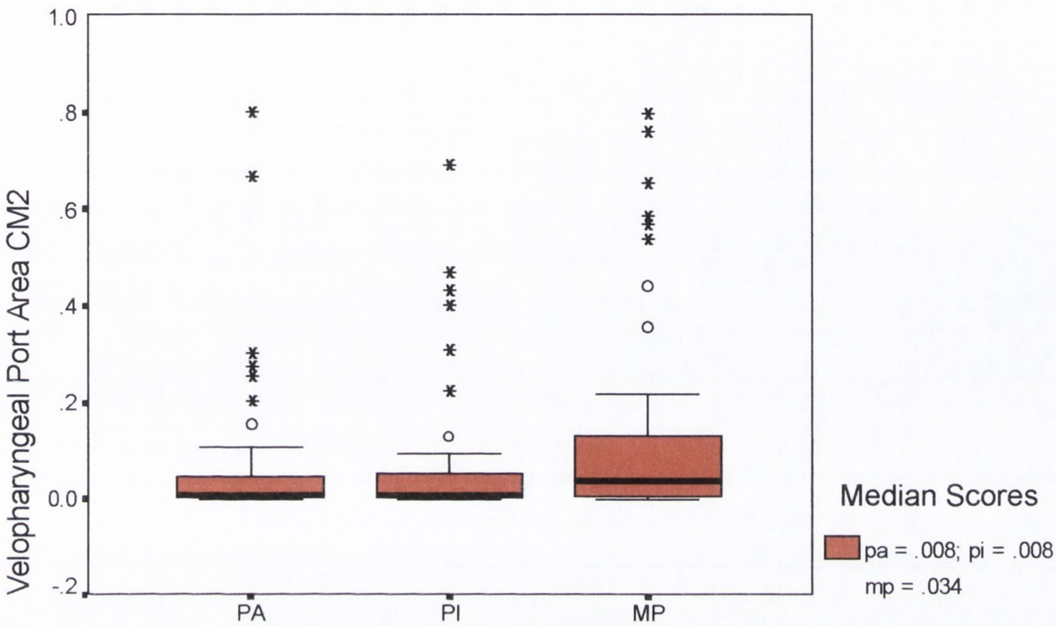


Figure 6.8 Distribution of velopharyngeal port size for the study group during the production of the speech samples. The graph indicates the median nasal pressure measure (thick black line) and the 50% of participants who fall between the 25th and 75th percentile (red box). The highest and lowest values (excluding outliers and extremes) is shown by thin black lines. The outliers (o) indicate values between 1.5 and 3 box-lengths from the upper edge of the box and the extremes (*) indicate values that are greater than 3 box-lengths from the upper edge of the box.

Independent sample t-tests indicated significant differences between velopharyngeal area measurements for the study group and for normal speakers during the production of all three speech stimuli ('pa' $t = 5.3$, $df = 199$, $p < .001$, 95% CI = 0.03,0.134; 'pi' $t = 5.9$, $df = 197$, $p < .001$, 95% CI = 0.03,0.119 and; 'hamper' $t = 7.6$, $df = 198$, $p < .001$, 95% CI= 0.074,0.196) (Figure 6.8).

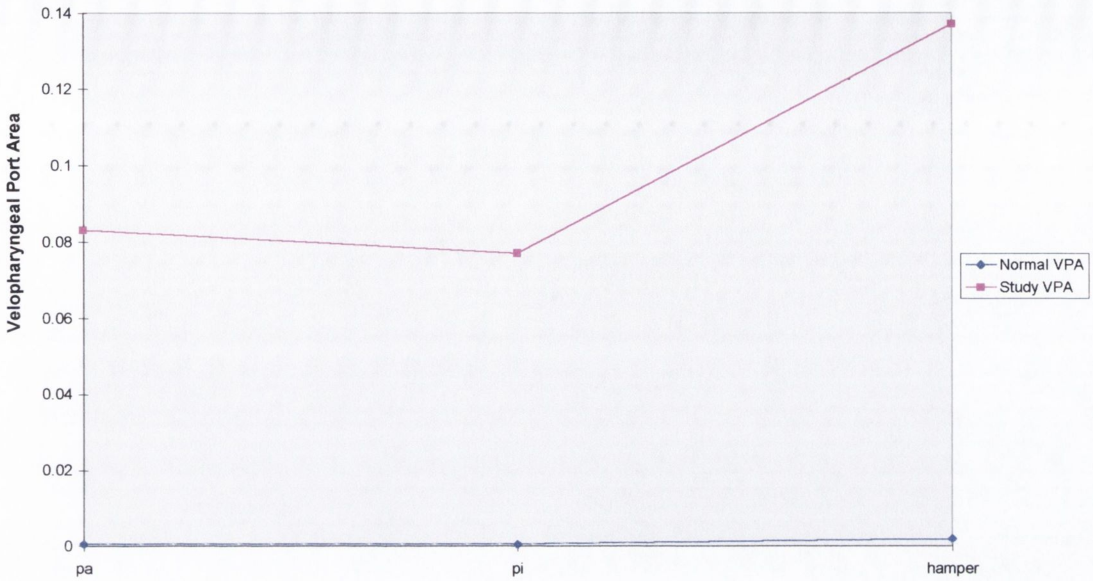


Figure 6.9 Mean velopharyngeal port area measurements for normal speakers and the study group during the production of the three speech stimuli.

6.4 Relationship between measurements

Pearson product moment correlations were carried out to evaluate the relationship between perceptual ratings of nasality and nasalance scores. Nonparametric Spearman-Rank order correlations were calculated to evaluate the relationship between perceptual ratings of nasal airflow errors and instrumental measurements. Where correlations were good, test sensitivity, test specificity and overall efficiency were also analysed. The relationships between the following measurements were analysed:

1. perceptual ratings of nasality and nasalance scores; and
2. perceptual ratings of nasal airflow and pressure/flow measures.

Correlational analyses between perceptual ratings of nasality and nasal airflow and instrumental measurements were carried out for each category of nasality (hypernasality and hyponasality) and each category of nasal airflow (nasal emission, nasal fricative, nasal turbulence and velopharyngeal fricative). Analysis was also carried out for the different speech stimuli used in instrumental assessments.

6. 4. 1 Relationship between Perceptual Measurements of Hypernasality and Nasalance Scores.

Results of correlational analyses indicated a substantial positive relationship between perceptual ratings of hypernasality and nasalance scores (Table 6.17). The best correlation was found between ratings of hypernasality and nasalance scores on the Total Test sentences and between ratings of hypernasality and nasalance scores on High Pressure consonant sentences. A low correlation was found between hypernasality rating and nasalance scores for the Nasal sentence. (Scattergram plots of relationships between perceptual ratings and nasalance scores are presented in Appendix 22).

Table 6.17 The Pearson correlation coefficients for the relationship between perceptual ratings of hypernasality and mean nasalance scores for each sentence category (n = 50).

Hypernasality Ratings and Nasalance Scores

	Total sentences	**HP sentences**	**LP sentences**	**Mixed sentences**	**Nasal sentence**
Correlation Coefficient	.74 $p < .001$	.74 $p < .001$	.69 $p < .001$	.72 $p < .001$	.46 $p < .01$

Test sensitivity and specificity, and overall efficiency indicated a good relationship between perceptual judgements of hypernasality and nasalance scores for the Total Test sentences, High Pressure consonant sentences and Low Pressure consonant sentences (Tables 6.18, 6.19, 6.20). For the Mixed consonant sentences, test sensitivity was moderate, while specificity was excellent (Table 6.21). Cut-off values were determined statistically for each sentence category using the normal mean score plus 2 standard deviations above the mean.

Table 6.18 Extent to which nasalance scores obtained during the production of the Total Test Sentences identified the presence or absence of hypernasality.

	Hypernasality		
Nasalance	Present	Absent	Total
#35	6	11	17
>35	30	3	33
Total	36	14	
	Sensitivity	Specificity	Overall Efficiency
	0.83	0.78	0.82

Table 6.19 Extent to which nasalance scores obtained during the production of High Pressure Consonant Sentences identified the presence or absence of hypernasality.

	Hypernasality		
Nasalance	Present	Absent	Total
#24	6	12	18
>24	30	2	32
Total	36	14	
	Sensitivity	Specificity	Overall Efficiency
	0.83	0.86	0.84

Table 6.20 Extent to which nasalance scores obtained during the production of Low Pressure Consonant Sentences identified the presence or absence of hypernasality.

	Hypernasality		
Nasalance	Present	Absent	Total
#28	4	11	15
>28	32	3	35
Total	36	14	
	Sensitivity	Specificity	Overall Efficiency
	0.88	0.78	0.86

Table 6.21 Extent to which nasalance scores obtained during the production of Mixed Consonant Sentences identified the presence or absence of hypernasality.

	Hypernasality		
Nasalance	Present	Absent	Total
#48	13	14	27
>48	23	0	23
Total	36	14	
	Sensitivity	Specificity	Overall Efficiency
	0.63	1	0.74

6. 4. 2 Relationship between Perceptual Ratings of Hyponasality and Nasalance Scores

Results indicated a moderate negative correlation between ratings of hyponasality and mean nasalance scores (Table 6.22). As expected, the highest correlation was found between nasalance scores for the nasal and mixed consonant sentences (containing nasal consonants) and hyponasality. The relationship was negative because as hyponasal ratings increase nasalance scores on a nasal sentence should decrease.

Table 6.22 Pearson product moment correlation coefficients for perceptual ratings of hyponasality and nasalance scores for the different sentence categories (n = 50).

	Total	HP consonant sentences	LP consonant sentences	Mixed consonant sentences	Nasal consonant sentence
Correlation Coefficient	-.67 $p < .001$	-.48 $p = .001$	-.56 $p = .001$	-.72 $p < .001$	-.76 $p < .001$

Analysis of test sensitivity and specificity indicated good test sensitivity and excellent test specificity and overall efficiency of the nasalance scores during the production of the Nasal sentence in identification of participants with hyponasality (Table 6.23).

Table 6.23 Extent to which nasalance scores obtained during the production of a Nasal Sentence identified the presence or absence of hyponasality.

	Hyponasality		
Nasalance	Present	Absent	Total
∃ 37	2	41	43
< 37	5	2	7
Total	7	43	
	Sensitivity	Specificity	Overall Efficiency
	0.71	0.95	0.92

6. 4. 3 Relationship between Perceptual Ratings of Nasal Airflow and Pressure/Flow Measurements

Spearman-Rank order correlations were calculated to assess the relationship between perceptual ratings of nasal airflow and pressure/flow measurements during the production of /p/ in 'pa', 'pi' and 'hamper'. Perceptual ratings of nasal airflow were ranked according to strength, frequency and consistency (Chapter 5, Section 6).

6. 4. 3 (i) Nasal Emission and Pressure/Flow Measurements.

Results indicated moderate statistically significant correlations for perceptual ratings of nasal emission and pressure/flow measurements on /pa/ and /pi/ ($p < .001$) (Table 6.24). Good significant correlations were found between perceptual ratings of nasal emission and three pressure/flow measurements (nasal flow, differential pressure and velopharyngeal port area) on 'hamper'. There was a weak, but significant correlation between perceptual ratings of nasal emission and oral pressure on 'hamper' ($p = .002$). Scattergrams of the relationship between nasal emission and pressure/flow measurements on /mp/ are presented in Appendix 23.

Table 6.24 Correlation coefficients for the relationship between perceptual ratings of nasal emission and pressure/flow measurements (nasal flow, oral pressure, differential pressure, velopharyngeal port area) for the three speech stimuli. * significant $p < .001$. ** $p < .01$

	'pa'				'pi'				'hamper'			
	nf	**op**	**dp**	**vpa**	**nf**	**op**	**dp**	**vpa**	**nf**	**op**	**dp**	**vpa**
Correlation Coefficient	.59	-.52	-.62	.57	.54	-.52	-.61	.55	.66	-.44	-.68	.67
Number	49	49	49	49	47	47	47	47	48	48	48	48
	*	*	*	*	*	*	*	*	*	**	*	*

Test Sensitivity & Specificity

Test sensitivity and specificity was carried out for instrumental measurements that had correlation coefficients approximating 0.6, indicating good relationships between perceptual measurements and instrumental measurements. As correlations were better for the production of 'pa' compared to 'pi', sensitivity and specificity were only analysed for measurements during 'pa' and 'hamper'. The following sensitivity and specificity analyses were carried out:

1. the extent to which nasal flow measurement indicated the presence or absence of nasal emission during the production of 'pa' and 'hamper' (Tables 6.25 & 6.26);
2. the extent to which differential pressure measurement indicated the presence or absence of nasal emission during the production of 'pa' and 'hamper' (Tables 6.27 & 6.28);
3. the extent to which velopharyngeal port area measurements indicated the presence or absence of nasal emission during the production of 'pa' and 'hamper' (Tables 6.29 & 6.30).

Test sensitivity, specificity and overall efficiency for nasal flow measurements were statistically determined. Cut-values were calculated statistically using the normal mean value + 2 standard deviations above the mean. A cut-off value of 4ml/s was used for analysis of nasal flow during the production of 'pa'. Results indicated excellent test sensitivity, good overall efficiency and weak test specificity for nasal flow measurements during 'pa' (Table 6.25).

Table 6.25 Extent to which nasal flow measurement during the production of 'pa' identified the presence or absence of nasal emission.

	Nasal Emission /pa/		
Nasal Flow	Present	Absent	Total
#4 ml/s	2	10	12
>4 ml/s	24	13	37
Total	26	23	
	Sensitivity	Specificity	Overall Efficiency
	0.92	0.43	0.69

A cut-off value of 14.5 ml/s was used for analysis of nasal flow during the production of /p/ in hamper. Analysis of test sensitivity, specificity and overall efficiency for nasal flow measurements on /p/ in 'hamper' indicated excellent sensitivity, good overall efficiency and moderate specificity (Table 6.26).

Table 6.26 Extent to which nasal flow measurements during the production of /p/ in 'hamper' identified the presence or absence of nasal emission.

	Nasal Emission /mp/		
Nasal Flow	Present	Absent	Total
#14.5 ml/s	1	13	14
>14.5 ml/s	25	9	34
Total	26	22	
	Sensitivity	Specificity	Overall Efficiency
	0.96	0.59	0.79

Test sensitivity and specificity for differential pressure measurements indicated good sensitivity, specificity and overall efficiency for /mp/, but poor sensitivity and overall efficiency and excellent specificity for /pa/ (Tables 6.27, 6.28). A cut-off value of 3 cm H_2O was selected, as normal differential pressures have been reported to be greater than or equal to 3 cm H_2O (Warren, 1979).

Table 6.27 Extent to which differential pressure measurements during the production of 'pa' identified the presence or absence of nasal emission.

	Nasal Emission /pa/		
Differential Pressure	Present	Absent	Total
$\geq$ 3 cm H_2O	15	22	37
< 3 cm H_2O	11	1	12
Total	26	23	
	Sensitivity	Specificity	Overall Efficiency
	0.57	0.95	0.66

Table 6.28 Extent to which differential pressure measurements during the production of /p/ in 'hamper' identified the presence or absence of nasal emission.

	Nasal Emission /mp/		
Differential Pressure	Present	Absent	Total
≥ 3 cm H_2O	7	18	25
< 3 cm H_2O	19	4	23
Total	26	22	
	Sensitivity	Specificity	Overall Efficiency
	0.73	0.81	0.76

Analysis of test sensitivity and specificity for velopharyngeal port area measurements in the identification of nasal emission indicated excellent test sensitivity, good overall efficiency and weak specificity (Table 6.29). A statistically determined cut-off value of .0017 cm^2 was used.

Table 6.29 Extent to which velopharyngeal port area measurements during the production of 'pa' identified the presence or absence of nasal emission.

	Nasal Emission /pa/		
Velopharyngeal Port Area	Present	Absent	Total
# .0017 cm^2	2	10	12
> .0017 cm^2	24	13	37
Total	26	23	
	Sensitivity	Specificity	Overall Efficiency
	0.92	0.43	0.69

Results indicated excellent sensitivity, moderate specificity and good overall efficiency of velopharyngeal port area measurements during the production of /p/ in 'hamper'(Table 6.30).

Table 6.30 Extent to which velopharyngeal port area measurements during the production of /p/ in 'hamper' identified the presence or absence of nasal emission.

Velopharyngeal Port Area	Nasal Emission /mp/ Present	Absent	Total
# .007 cm^2	1	13	14
> .007 cm^2	25	9	34
Total	26	22	
	Sensitivity	Specificity	Overall Efficiency
	0.96	0.59	0.79

6. 4. 3 (ii) Nasal Fricative and Pressure/Flow Measurements

Poor correlations were found between perceptual ratings of nasal fricative and pressure/flow measurements (Table 6.31). Correlation coefficients ranged from -.02 to .08. None were significant. Test sensitivity and specificity were not carried out due to weak correlations.

Table 6.31 Spearman-Rank correlation coefficients for the relationship between perceptual ratings of nasal fricative and pressure/flow measurements (nasal flow, oral pressure, differential pressure, velopharyngeal port area (vpa)) during the production of 'pa', 'pi' and 'hamper'.

	'pa'				**'pi'**				**'hamper'**			
	nf	**op**	**dp**	**vpa**	**nf**	**op**	**dp**	**vpa**	**nf**	**op**	**dp**	**vpa**
Correlation Coefficient	.02	-.02	-.07	-.05	.02	-.13	-.12	-.02	.08	-.08	.04	-.12
Number	49	49	49	49	47	47	47	47	48	48	48	48

6. 4. 3 (iii) Nasal Turbulence and Pressure/Flow Measurements

Moderate positive significant correlations found between perceptual ratings of nasal turbulence and oral pressure measurements for 'pi' ($r = .51$, $p < .001$) and for 'pa' ($r = .42$ $p < .01$) and between perceptual ratings of nasal turbulence and differential pressure for 'pi' ($r = .46$, $p = .001$) (Table 6.32). All other correlations were poor and non significant. Test sensitivity and specificity were not analysed due to weak correlations.

Table 6.32 Spearman-Rank correlation coefficients for the relationship between perceptual ratings of nasal turbulence and pressure/flow measurements (nasal flow, oral pressure, differential pressure, velopharyngeal port area) during the production of 'pa', 'pi' and 'hamper'.
* significant $p < .001$

	'pa'				'pi'				'hamper'			
	nf	**op***	**dp**	**vpa**	**nf**	**op***	**dp***	**vpa**	**nf**	**op**	**dp**	**vpa**
Correlation Coefficient	-.06	.42	.39	-.07	-.14	.51	.46	-.18	-.19	.22	.24	-.21
Number	49	49	49	49	47	47	47	47	48	48	48	48

6. 4. 3 (iv) Velopharyngeal Fricatives and Pressure/Flow Measurements

Correlations between velopharyngeal fricative and pressure/flow measures were poor. Nasal flow and velopharyngeal port area had weak but significant correlations with perceptual ratings of velopharyngeal fricatives (Table 6.33). Test sensitivity and specificity were not analysed due to weak correlations.

Table 6.33 Spearman-Rank correlation coefficients for the relationship between perceptual ratings of velopharyngeal fricatives and pressure/flow measurements (nasal flow, oral pressure, differential pressure, velopharyngeal port area) during the production of 'pa', 'pi' and 'hamper'. * significant $p < .01$

	'pa'				**'pi'**				**'hamper'**			
	nf	**op**	**dp**	**vpa**	**nf**	**op**	**dp**	**vpa**	**nf**	**op**	**dp**	**vpa**
Correlation Coefficient	-.40	.10	.10	-.37	-.42	.21	.24	-.43	.39	-.26	.32	-.36
Number	49	49	49	49	47	47	47	47	48	48	48	48
	*			*	*			*	*			*

CHAPTER 7

DISCUSSION

Perceptual judgements should form the basis of any clinical assessment of speech, as a speech disorder does not exist, until it is perceived by the listener (Kuehn, 1982). However, the reliability of perceptual judgements of speech has been questioned and problems in evaluating and reporting reliability have been demonstrated (Kearns & Simmons, 1988; Wirz & Mackenzie Beck, 1995). The importance of reporting the percent of agreement and a reliability measurement, which takes account of chance agreements, has been highlighted (Kreiman et al., 1993; Watterson et al., 1998b). Intra- and inter-rater reliability of judgements of nasality and nasal airflow errors were evaluated, using the perceptual scale. These reliability studies have been reported in Chapter 4. In the development of the Perceptual Profile the known factors which might influence reliability and agreement were taken into account. For example, the phonetic context of the speech sample has been carefully developed to allow for assessment of phonemes which are vulnerable to velopharyngeal dysfunction (Sell et al., 1994). Vocal quality and articulatory errors were recorded as possible influencing factors. Listening conditions and the use of anchor stimuli were evaluated (Chapter 4, Sections 4 & 6). The following discussion presents a brief description of the study population. The results of the perceptual assessment of the study group are discussed in section 7. 2.

Studies on the use of nasometry have indicated the need for normative data for varying languages and dialects (Seaver et al., 1991; Trindade et al., 1997; van Doorn & Purcell, 1998). As a result, normative data was obtained from normal English-speaking Irish children and results have been presented in Chapter 3. The following discussion compares results of the study group with results of the normal Irish-speakers. The validity of the Nasometer is evaluated by comparing the relationship between nasometry scores and perceptual ratings of nasality in the study population. The varying cut-off values are evaluated. Karnell (1995) hypothesized that the use of separate HP (HP) consonant and Low Pressure (LP) consonant speech stimuli may help identify participants with nasal turbulence. This hypothesis is discussed in the light of the present results.

Pressure/flow measurements using the PERCI SARS are compared with pressure/flow measurements from normal children of similar age (normal data reported by Zajac, 1998, personal communication). No previous studies have reported on the relationship between perceptual measures of nasal airflow problems and pressure/flow measurements. In the present study the main focus of the pressure/flow discussion is on this relationship.

The relationships between perceptual and instrumental measurements were analyzed using correlational analysis, test sensitivity and specificity studies. Correlational analyses were carried out in order to ascertain if the perceptual assessments rated the speech in a similar manner to the instrumental measurements. Test sensitivity is the percentage of participants that are identified as having abnormal speech on one test and who are also identified as having abnormal speech on another test. Test specificity refers to the percentage of patients who are identified as having normal speech on one test and who are also identified as having normal speech on another test. An overall efficiency rating can be calculated adding the number of times one test is in agreement with the other test, divided by the total number of opportunities (i.e. total number of participants tested) (Watterson et al., 1998a).

Cut-off values for test scores have been determined to distinguish normal and abnormal speakers. These values can be determined statistically, by using 2 standard deviations above or below the mean for a normal population. In the present study, statistically determined cut-off values were used to evaluate sensitivity and specificity. Cut-off values can also be determined clinically, by analysing results from a clinical population of pathological speakers to determine what test score best identifies the presence and absence of the speech error. Clinically determined cut-off values were evaluated in the present study and compared to statistically determined cut-off values in order to ascertain an optimum cut-off value for test scores.

Possible causes of discrepancies in perceptual ratings and instrumental measurements are discussed in section 7. 6 and section 7.13. The efficiency of the speech stimuli used for the instrumental assessments is discussed.

In order to evaluate these aspects of the instrumental techniques, each technique is discussed under the following headings:

- efficiency of the instrument to differentiate between normal and pathological speakers
- relationship between perceptual ratings and instrumental measurements
- factors affecting the relationship between perceptual ratings and instrumental measurements
- efficiency of cut-off values
- speech stimuli used for the instrumental assessments

Limitations of the study are presented in section 7. 16. In section 7.17, the findings of the present study are summarised, while the conclusions are presented in Chapter 8.

7. 1 Study Population

In the present study, the age of the study group ranged from 4;10 years to 15;10 years. In previous studies of comparison of instrumental and perceptual measurements, both adults and children were included (Dalston & Warren, 1986; Laine et al., 1988; Dalston et al., 1991a; Dalston et al., 1991b; Warren et al., 1994). However, some of these studies presented separate results for the two age groups, thus facilitating comparison with the present results (Dalston et al., 1991a; Warren et al., 1994). Studies comparing nasalance scores and perceptual measurements of nasality for children with a similar age range to the present study, were carried out by Watterson et al. (1993), and Watterson et al. (1996). Hence, the populations in these studies are more directly comparable to the present study than other populations. In the present investigation, the study group consisted of a consecutive series of children referred to the National Cleft Palate Unit for investigation of nasality and nasal airflow problems, with a wide range of aetiologies. Diagnosis of the 50 participants in the study group indicated that almost half the group (48%) had velopharyngeal dysfunction in the absence of a history of overt cleft palate. In contrast, previous studies had a smaller percentage of participants with velopharyngeal dysfunction without cleft palate (20%) (Dalston et al., 1991a; Dalston et al., 1991b; Watterson et al., 1993).

7. 2 Perceptual Assessment

The problems of perceptual assessments of nasality and nasal airflow were identified in the review of literature (Chapter 2). Kent (1996) reported that raters did not use equivalent definition of terms, and failed to reach consensus on which scale should be used. Wirz and Mackenzie Beck (1995) reported that many of the scales used in speech assessment had no descriptors for points on the scale. As a result, the reproducability and inter-rater reliability could not be guaranteed. The Perceptual Profile used in the present study attempts to overcome some of these problems. The perceptual scale is descriptive rather than numerical in nature, with each point on the scale being defined in terms of what is perceived by the listener. Other studies have used a numerical scale where the individual points on the scale were not defined and only the first and last points on the scale were described in terms of mild or severe (Dalston & Seaver, 1992; Watterson et al., 1993).

The nasal airflow section of the scale is different to the nasality section, in so far as different categories of nasal airflow are defined and rated according to strength, frequency and consistency. The descriptive scale of nasal airflow is unlike scales used in previous studies, where numerical scales of nasal emission and/or turbulence were used (Sell et al., 1994).

Perceptual judgements have been described as the "gold standard" against which other assessment techniques can be compared (Dalston & Warren, 1986). But, using perceptual judgements as the "gold standard" has been shown to have inherent problems (Kuehn, 1982; Kent, 1996). The reliability of perceptual assessments of nasality and nasal airflow errors is influenced by factors such as phonetic context, voice and articulatory errors, listening conditions and listener training (Chapter 2, Section 5.2). These factors were considered in the development and use of the Perceptual Profile developed for this study. The phonetic context of the perceptual assessment and Nasometry were carefully selected, while a live listening situation was used for the perceptual assessment. Articulatory and voice errors were recorded as factors which may influence ratings. The rater in the present study was highly familiar with and trained in the use of the Perceptual Profile. The reliability studies indicated that there was excellent agreement and overall good reliability for the author in the use of the Perceptual Profile. As a result, it is reasonable to compare the author's ratings with instrumental measurements. However, despite the careful consideration of factors influencing the perception of nasality and nasal airflow errors in speech, the inter-rater reliability was only adequate for some speech parameters. Results indicated acceptable levels of agreement between raters compared to previous studies (Hardin et al., 1992; Lohmander-Agerskov, 1996; Watterson et al., 1998a), but there was discrepancy between agreement and kappa scores in the present study, which indicates the fact that high levels of agreement do not always imply good reliability. The varying inter-rater reliability and the fact that correlation and agreement were used in previous studies, questions the view that perceptual assessment is the "gold standard". Perhaps it is better to refer to the perceptual assessment as a baseline assessment against which instrumental measurements can be compared. The relationship between the Perceptual Profile and instrumental assessments is discussed under the sections reviewing nasometry and pressure/flow measurements.

Conclusions from the reliability studies of the Perceptual Profile (Chapter 4), indicate that this assessment has good intra-rater reliability for the rater in the main study. Inter-rater reliability varied across raters and speech parameters. The overall results underscore the need to supplement perceptual assessments with instrumental measurements.

7. 2. 1 Perceptual Assessment of Nasality

Assessment results indicated that 20% of participants had no perceptible hypernasality, but had nasal airflow problems only. Forty two percent of the participants had mild nasality problems indicative of borderline velopharyngeal dysfunction (Morris, 1984). This mild group, according to Morris (1984), can be further divided into two groups according to the characteristics of velopharyngeal function: almost-but-not-quite-adequate (mild consistent hypernasality) and sometimes-but-not-always (mild inconsistent hypernasality). Management decisions for these groups should be undertaken with caution and following in-depth and thoughtful assessment (D'Antonio & Scherer, 1995).

Interestingly, 67% of participants who were perceived as hypernasal, had inconsistent nasality. Assessments of consistency of nasality, using this profile, were found to have good intra-rater reliability (a kappa score of 0.7 and 87% agreement). However, inter-rater agreements on consistency of nasality ranged from good (a kappa score of 0.7 and 85% agreements) to poor (a kappa score of 0.1 and 81% agreement). This weak inter-rater reliability may have been due to varying interpretations of the term 'consistency' by the raters. Although there was questionable reliability of assessment of consistency of nasality, this speech parameter cannot be ignored, as consistency may influence the relationship between perceptual and instrumental measurements, and appears to be of clinical significance, according to Morris (1984). This will be discussed in further detail in sections 7.6 and 7.13.

Only 14% of the population studied had perceptible hyponasality. Hence, results of instrumental assessments of hyponasality should be considered with caution.

7. 2. 2 Perceptual Assessments of Nasal Airflow

Over half the population studied presented with nasal emission. However, of these, 81% had weak nasal emission. Approximately 44% had mild consistent errors. Eight participants (16%) presented with a nasal fricative, of which seven had phoneme specific errors. Thirty four percent of the study group had perceptible nasal turbulence. All participants presenting with nasal turbulence fell into the mild category. Only 14% (n = 7) of the study group had a velopharyngeal fricative. All of the velopharyngeal fricatives, with the exception of one, were phoneme specific. Of the six participants with phoneme specific velopharyngeal fricatives, two participants had no other nasality or nasal airflow problems. These participants had what has been referred to in the literature as " phoneme specific nasality" (Peterson, 1975; Trost, 1981; Albery, 1989). This type of problem is usually indicative of velopharyngeal mislearning, as no structural abnormality is detected.

Analysis of perceptual results indicated that over half the population studied had mild nasality and/or nasal airflow problems. These types of problems have been found to be difficult to assess reliably (McWilliams & Philips, 1979). The earlier reliability study indicated that when nasal airflow errors were at the mild end of the scale, there was confusion in categorisation of errors. It has also been found that, when participants present with mild speech errors, there is a weaker relationship between perceptual and instrumental measurements (Warren, 1998, personal communication). This will be discussed further in the analysis of instrumental results.

In summary, the population assessed in this study was typical of the caseload referred for speech investigation to the National Cleft Palate Unit. However, this population may produce lower than expected relationships between measurements, due to the large number of participants with mild speech errors. Nevertheless, it is representative of what is generally considered to be the most difficult group to evaluate.

7. 3 Nasometry

Section 7.3.1 of this discussion compares nasalance results from the two populations: a group of children with nasality/nasal airflow problems and a group of normal speakers of similar age, language and dialect. In order to compare the present results with the literature, the nasalance scores are presented for each category of speech stimuli. The Mixed consonant sentence category was not intended as a diagnostic category, but was included to make up the Total Test sentence category. As a result, the nasalance scores for the Mixed consonant category will not be discussed. The relationship between perceptual ratings of nasality and nasalance scores are discussed in section 7.4. The value of various cut-off nasalance scores in identifying nasality problems is discussed in section 7.5. Factors that may have influenced the relationship between perceptual ratings and nasalance scores are discussed in section 7.6. Finally, section 7.7 discusses the benefit of using separate HP consonant and LP consonant stimuli on the basis of present results.

7. 3. 1 Effectiveness of the Nasometer in the Discrimination of Normal and Pathological Speakers

In the present study two distinct populations were used for comparison of nasalance scores. The normal population had no history of cleft palate or velopharyngeal dysfunction, while the study population had such a history. The study group consisted of participants with perceived nasality and/or nasal airflow problems. In order to compare nasalance scores of the present study group with nasalance scores of normal speakers, data from the nasalance study for normal English-speaking Irish children was used (Chapter 3). Where possible, the present results were compared to results from other studies in the literature. However, populations studied in the literature differed greatly in terms of presenting problems, ages, number of participants, languages and dialect. In most of these studies the population consisted of participants who were under review at a cleft palate/craniofacial centre, some of whom were reported to have normal speech, and others having speech problems related to cleft palate or velopharyngeal dysfunction.

Results indicated that there was a statistically significant difference between the mean nasalance scores for the study group and mean nasalance scores for normal speakers across all speech stimuli ($F = 69.5$, $p < .001$). A post-hoc analysis indicated significant differences in the mean nasalance scores between the two groups of speakers during production of all the speech stimuli. Hence, the Nasometer was able to distinguish between the group of normal speakers and the group of speakers with nasality and nasal airflow errors.

The mean nasalance scores of the pathological speakers in the present study compared well with mean nasalance scores for similar populations in other studies (Watterson et al., 1993; Watterson et al., 1996; Watterson et al., 1998a) (Table 1). However, only one study reported nasalance scores for LP consonant sentences in a clinical population (Watterson et al., 1998a). They found a lower mean nasalance score of 29% for LP consonant sentences compared to 38% in the present study. One possible explanation for this discrepancy is that there were only 25 participants in the Watterson et al. (1998a) study compared to 50 participants in the present study. The larger study group resulted in a wider range of nasalance scores and a higher mean score.

Table 7.1 Mean nasalance scores for different study groups during the production of comparable speech stimuli.

Speech Stimuli	**Present study**	**Watterson et al. (1993)**	**Watterson et al. (1996)**	**Watterson et al. (1998)**
Total Test Sentences	40%	44%	43%	
HP consonant Sentences	33%	31%	30%	30%
LP consonant Sentences	38%			29%

A significant, but small, difference was found between nasalance scores for the study group and nasalance scores for normal speakers during the production of a nasal sentence ($p = .009$). Previous studies have reported that sentences containing a high proportion of nasal consonants were not useful in identifying patients with hypernasality (Dalston et al., 1991b; Nellis et al., 1992). Although the difference between the mean nasalance scores for the two groups was significant, the difference was small (2.37%) and the confidence interval was close to zero (95% CI = .01,4.73). Although small differences in nasalance scores between groups of speakers have been found to be statistically significant in other studies (Seaver et al., 1991; Karnell, 1995), the clinical significance of these findings has been questioned (Trindade et al., 1997; Watterson et al., 1998a). The lack of clinical significance can be explained by the individual variation of the nasalance scores during the production of the Nasal sentence in the present study (5% or more). This variation has also been reported in other studies (Seaver et al., 1991; van Doorn & Purcell, 1998). Therefore, although the small difference of 2% may be significant for a large group of speakers, it is not clinically significant for individual speakers. Hence, these findings do not support the use of the Nasal sentence for classification of speakers as hypernasal or normal.

The mean standard deviation scores and the range of nasalance scores for all speech stimuli reported in the present study differed from mean standard deviation scores and ranges in previous studies (Watterson et al., 1996: Watterson et al., 1998a). There were some methodological differences between the studies. In the Watterson et al. (1996;1998a) studies, a smaller number of participants were assessed (20 and 25 respectively). This small number of participants in the study and the possible differences in presenting problems of nasality may explain the discrepancy in the standard deviations and range of scores in the three studies.

7. 4 Relationships Between Nasalance Scores and Perceptual Ratings of Nasality

For the purpose of statistical analysis of the relationship between perceptual ratings of nasality and nasalance scores, it can be argued that the descriptive scale is an interval scale, in that the difference between a rating of mild/acceptable and mild/moderate is equal to the difference between mild/moderate and moderate. In considering the descriptive scale as an interval scale, the Pearson product moment correlation can be used to compare perceptual ratings of nasality with nasalance scores. This also allows for comparison with previous studies. Dalston and Seaver (1992) argued that nasality scales can be considered to be interval or ordinal scales. In the present study, the nasality scale was considered as an interval scale and a Pearson correlation coefficient was calculated. The Spearman Rank order correlation was carried out for the comparison of perceptual ratings of nasal airflow errors with instrumental measurements, as the nasal airflow scales were categorical. That is, nasal airflow errors were categorized as weak or strong, consistent or inconsistent, frequent or infrequent on the Perceptual Profile. However, these categories are not in a numerical order where 1 is always less severe then 2 and 5 is always more severe than 4.

Correlation coefficients between perceptual ratings of nasality and nasalance scores are presented in the present study in order to compare results with previous studies. Test sensitivity and specificity provide an overall efficiency score that indicates the general ability of the nasometer to distinguish between normal and abnormal speakers. Cut-off values were statistically determined by using two standard deviations above the mean nasalance scores for normal speakers. Both types of analyses (correlation and test sensitivity, test specificity and overall efficiency) will be discussed to evaluate the relationship between perceptual ratings of nasality and nasalance scores for each sentence category.

7. 4. 1 The Relationship between Nasalance Scores for the Total Test Sentences and Perceptual Ratings of Nasality

A statistically significant correlation was found between perceptual ratings of nasality and nasalance scores for the Total Test sentences ($r = .74$ $p < .001$). This correlation indicated a stronger relationship between perceptual ratings of nasality and nasalance scores for speech stimuli containing all consonant types compared with previous studies. Paynter et al. (1991) and Dalston and Seaver (1992) reported a Pearson correlation coefficient of 0.63. There are several possible explanations for the differences in results. Firstly, different speech stimuli were used for the perceptual assessment and for the nasometric assessment in the study by Dalston and Seaver (1992). A conversational speech sample was used for the perceptual assessment, while the Rainbow Passage (phonetically equivalent to the Total Test sentences in the present study) was used for nasometric assessment. In contrast, the speech sample used for the perceptual assessment in the present study included the Total Test sentences. In the Paynter et al. (1991) study, there was no difference in the speech stimuli used in each assessment. The speech stimulus used for the nasometric assessment was audio recorded and later analyzed perceptually by a panel of judges. The difference in results between the present study and the Paynter et al. study may be due to the fact that in the latter study a panel of judges was used to rate nasality perceptually. Watterson et al. (1998b) reported that when panels of listeners are used to obtain nasality ratings, important individual variability of ratings is lost and this has been found to result in weak correlations between perceptual ratings of nasality and nasalance scores.

Another explanation for different correlations is that not all points on the perceptual scales used in the two previous studies were defined. The lack of scalar definition, according to Wirz and Mackenzie Beck (1995), may result in poor reliability. Dalston and Seaver (1992) did not report on the intra-rater reliability of the listener judgments. Intra-rater reliability for nasality judgments on the Perceptual Profile used in the present study was good (89% agreement and kappa = 0.7). It is possible that the use of descriptors instead of scalar points on the present scale may have improved reliability and hence, produced a better correlation between nasalance and nasality ratings. The present findings compare very favorably with an earlier study examining the relationship between Tonar II (the early version of the Nasometer) and listener judgments. Dalston and Warren (1986) found a correlation coefficient between ratings of nasality and nasalance scores of 0.76. In this study Dalston and Warren defined the five points on the rating scale and reported inter-rater reliability as ranging from 0.73 to 0.81. Hence, it appears that by using reliable judgments of nasality and defining all the points on the perceptual scale of nasality, the correlation between perceptual ratings and nasalance can be improved.

In contrast, two studies carried out by Watterson et al. (1993) and Watterson et al. (1996) reported poor non-significant correlations between nasalance scores for passages containing all consonant types and listener judgments of nasality. Watterson et al. (1993) accounted for the differences between their results and Dalston and Seaver's (1992) results by the use of only one listener for the perceptual analysis and the lack of demonstrated reliability in Dalston and Seavers's study. In the Watterson et al. (1993; 1996) studies a panel of listeners was used to rate nasality, and reliability was demonstrated using correlational analysis. However, more recently Watterson et al. (1998b) stated that using a panel of listeners do not necessarily indicate that the perceptual assessment is reliable. They pointed out that correlational analyses indicate that raters rate in a similar manner, but does not report agreement. Hence, Watterson et al.'s earlier speech results should be considered with caution. In the present study intra-rater reliability was good, while inter-rater agreement was acceptable. Thus, the reliability of the perceptual assessment of nasality and the methodology used in the present study may have resulted in improved correlations between perceptual ratings of nasality and nasalance scores for a speech stimulus containing all type of consonants.

Correlation, sensitivity and specificity results indicate that there was a good relationship between nasalance scores for the Total Test Sentences and perceptual ratings of nasality using the descriptive scale.

A cut-off value of 35% was chosen for the Total Test Sentences indicating 2 standard deviations above the mean. Results indicate that 83% of cases who were rated as hypernasal on the perceptual scale had nasalance scores above the cut-off value of 35% and 78% of cases who were rated a normal had nasalance scores below 35%. The overall efficiency (i.e. the sum of the number of times the Nasometer agreed with the perceptual judgments divided by the total number of opportunities) was 0.82. It would appear therefore that, when using the Total Test Sentences speech sample on the Nasometer, the number of subjects judged perceptually to be hypernasal compared well to the number of subjects identified as abnormal on the Nasometer. The number of subjects judged perceptually to be normal compared reasonably well to the number identified as normal by Nasometer.

No previous studies assessed sensitivity and specificity of the Nasometer for sentences containing all types of consonants. Indeed, previous studies reported that the Rainbow Passage (equivalent to the Total Test sentences) was not a good indicator of hypernasality. The present findings show that the use of a speech sample containing all consonant types with the Nasometer and the perceptual nasality scale developed for this study are reliable and valid assessment techniques.

7. 4. 2 The Relationship between Nasalance Scores on High Pressure Consonant Sentences and Perceptual Ratings of Nasality

Good correlations were found between nasalance scores for HP consonant sentences and hypernasality ($r = .74$, $p < .001$). Previous studies reported correlation coefficients ranging from 0.49 (Watterson et al., 1993) to 0.66 (Paynter et al., 1991). However, Watterson et al. (1998a) reported a slightly higher correlation of 0.78. This improved correlation from their earlier study may be explained by the fact that Watterson et al. (1998a) rated the HP consonant sentences separately from the LP consonant sentences. The good correlations between HP consonant sentences and perceptual ratings of nasality found in the present study may be due to the use of the descriptive scale of nasality and the good reliability of the scale.

Test sensitivity and specificity were assessed for nasalance scores using HP consonant sentences and a cut-off value of 24%. Results indicate good correspondence between nasalance scores and perceptual judgments of nasality. There was no difference between the sensitivity scores for the Total Test sentences and the score for HP consonant sentences. However, specificity for the HP consonant sentences increased to 0.86 compared to a specificity of 0.78 for the Total Test sample. This would indicate that a speech sample containing HP consonants and devoid of nasal consonants has a better correspondence with perceptual judgments of nasality. This supports the results of Dalston et al. (1991b) and Watterson et al. (1996) who report that a passage devoid of nasal consonants was a good predictor of velopharyngeal dysfunction.

The present results compare well with two previous studies. Using a cut-off value of 26%, Watterson et al. (1998a) reported a sensitivity of 0.84, and specificity of 0.88. Hardin et al. (1992) reported a sensitivity of 0.76, specificity of 0.85 and overall efficiency of 0.82 for the Zoo Passage using a cut-off value of 26%. Dalston et al. (1991b) reported a sensitivity score of 0.89 and a specificity of 0.95 and overall efficiency of 0.93, using a cut-off value of 32%. However, in Dalston et al.'s (1991b) study, only one judge was used for the perceptual judgments and no inta-rater reliability was reported. Furthermore, Dalston et al. (1991b) used a clinically determined cut-off value to assess sensitivity and specificity, whereas a statistically determined cut-off value was used in the present study. When a clinically determined cut-off value was adopted for the HP consonant sentences, this yielded no improvement in the overall efficiency in the present study (see Section 7.5).

7. 4. 3 The Relationship between Nasalance Scores on Low Pressure Consonant Sentences and Perceptual ratings of Nasality

The relationship between the nasalance scores for LP consonant sentences and perceptual ratings of nasality ($r = 0.69$) was weaker in the present study than the relationship found in the Watterson et al. (1998a) study ($r = 0.77$). The stronger relationship in the Watterson et al. (1998a) study may be explained by the methodological difference in the two studies. In the Watterson et al. (1998a) study, two separate perceptual ratings were made; one for HP consonant sentences and one for LP consonant sentences. Correlations were then calculated to ascertain the relationship between the perceptual ratings for the HP consonant sentences and nasalance scores for the same sentences. Further correlations were calculated for perceptual ratings on LP consonant sentences and nasalance scores on LP consonant sentences. The use of identical speech stimuli for both perceptual and nasometric assessment may have resulted in the stronger relationship between measurements.

In the present study test sensitivity indicated good correspondence between the perceptual judgment of the nasality and nasalance scores using LP consonant sentences. Specificity results indicated a moderate relationship between perceptual judgments and nasalance scores. A good overall efficiency was found (0.86). The test sensitivity scores and overall efficiency compare well with the Watterson et al. (1998a) study. However, test specificity was lower in the present study, indicating that normal speakers were misclassified as hypernasal in 31% of cases. The correlation results and lower specificity indicate that LP consonant sentences are not good indicators of hypernasality.

7. 4. 4 The Relationship between Nasalance Scores on the Nasal Sentence and Perceptual Rating of Hyponasality

A good significant correlation was found between nasalance scores during the production of the nasal sentence and perceptual ratings of hyponasality (r = -0.76). This relationship was stronger than the relationship reported by Dalston et al. (1991a) (r = -0.68). The present result was encouraging in the light of the poor reliability of the Nasometer when a limited speech stimulus was used to obtain the nasalance score (Nichols, 1999). However, Nichols (1999) pointed out in a study with many participants, equivalent to the present study, the overall nasalance means may be quite reliable, whereas individual nasalance scores may not. In contrast, Watterson et al. (1999) reported good correlations between nasalance scores for long and short utterances, indicating good validity for a limited speech sample. Hence, although the reliability of the nasalance scores for the nasal sentence was poor in the present study, the validity was good. This may be due to the descriptive nature of the perceptual scale.

In the present study, test sensitivity and specificity compared nasalance scores on the Nasal sentence with perceptual ratings of hyponasality. The cut-off score for hyponasality was 37%, based on two standard deviations below the mean for normal production of the nasal sentence (mean 51%, standard deviation 7%). Results indicated a test sensitivity of 0.71, test specificity of 0.95 and an overall efficiency of 0.92. These results are considerably better than results reported by Dalston et al. (1991a), (test sensitivity of 0.48 and specificity of 0.79).

Dalston et al. (1991a) reported that when patients with audible nasal emission were eliminated from the analysis, the relationship between nasometry scores and hyponasality changed dramatically. Sensitivity was reported as 1.0, specificity was reported as 0.85, and overall efficiency was found to be 0.90. Hardin et al. (1992) reported similar findings for a similar population, using a cut-off value of 50%. Participants in the present study, who were hyponasal, had no audible nasal emission and would therefore be similar to the groups studied by Dalston et al. (1991b) and Hardin et al. (1992). However, with similar populations, sensitivity was lower in the present study. The difference in results may be due to the limited speech sample used in the present study for assessment of hyponasality.

7. 4. 5 Summary

Correlational analyses, test sensitivity and specificity indicated that there was a good relationship between listener ratings of nasality and nasalance scores. The best relationship was found between nasalance scores for HP consonant sentences and perceptual ratings of hypernasality, and between nasalance scores for the Total Test sentences and perceptual ratings of hypernasality. The relationship between the nasalance scores for the Nasal sentence and perceptual ratings of hyponasality varied from acceptable to good. In this study, the findings indicated that the speech stimuli used with the Nasometer and the Perceptual Profile provided valid measurements of nasality.

7. 5 Efficiency of Cut-off Values

Test sensitivity and specificity analyses used statistically determined cut-off values (i.e. two standard deviations above the mean for normal speakers) (van Doorn & Purcell, 1998). In order to obtain the clinically determined cut-off value (i.e. the highest overall efficiency rating), varying cut-off values were used and overall efficiency was calculated. Examination of the data indicated that changes of 5% above or below the statistically determined cut-off value resulted in changes in overall efficiency.

Varying results were obtained from each sentence category.

- In the Total Test sentences the value of 35% (i.e. two standard deviations above the mean) yielded overall efficiency of 0.82, whereas a cut-off value of 30% indicated an improved overall efficiency rating of 0.88.
- For High-Pressure consonant sentences there was no difference in overall efficiency rating for values 5% above and below 24%.
- Improved overall efficiency rating was found for the LP sentences when a cut-off value of 23% was used. This cut-off provided an efficiency rating of 0.88, compared to an overall efficiency rating of 0.86 using a cut-off of 28%.

- Using a cut-off value of 42% for the Nasal sentence, there was limited difference in the overall efficiency.

The results would indicate that a range of cut-off values may best help identification of participants with abnormal nasality. For example, a nasalance score on the Total Test Sentence sample between 30% and 35% should be considered as a possible indication of hypernasality, but additional data would be required for this participant. Van Doorn & Purcell (1998; p. 291) stated that cut-off scores should 'act as a guide only to the limits of nasalance that correspond to the perception of normal resonance'.

7. 6 Factors which influence the Relationships between Perceptual Ratings of Nasality and Nasalance Scores.

Two important factors were expected to influence the relationship between perceptual and nasometry measurements. These were consistency of the speech error and associated nasal airflow errors. Consistency of hypernasality has recently been included in perceptual analyses of nasality (Sell et al., 1999). One would expect that inconsistent hypernasality would influence the correlation between hypernasality and nasometry. If the degree of nasality's variable, it is possible that the degree of nasality will be different during the nasometric test and the perceptual assessment. In the present study, the perceptual evaluation was carried out prior to nasometric and pressure/flow testing, although all assessments were carried out at the one assessment session. Results indicated a linear relationship and good correlations between nasalance scores for the Total Test Sentences and perceptual judgments of nasality (Appendix 22, Scattergrams). However, twelve participants were identified on the scattergram as having a weak relationship between measurements. Detailed analysis indicated that nine of the twelve participants who had a discrepancy between perceptual and nasometric measurements were perceived to have inconsistent hypernasality.

All of the 12 participants also had perceived nasal airflow problems and increased pressure/flow measurements. Karnell (1995) hypothesised that audible turbulent nasal airflow on HP consonant sentences may be detected as nasal acoustic energy by the Nasometer. Hence, it was anticipated that if nasal airflow problems resulted in increased nasalance scores, then nasalance scores for HP consonant sentences should have been greater than nasalance scores for LP consonant sentences, as nasal airflow errors are unlikely to occur on LP consonants. This was not the case (see Section 7.7). Correlation coefficients did not change when participants with nasal airflow errors were excluded from the analyses. However, it has been reported that nasal airflow errors in speech can influence the perception of hypernasality (McWilliams et al., 1990) and this may have influenced the ratings on the perceptual assessment.

The present results indicated that, although there was a linear relationship between perceptual ratings of nasality and nasalance scores, high nasalance did not always mean high perceptual rating of hypernasality. Watterson et al. (1993) stated that lack of agreements between the Nasometer and perceptual meetings may be partly due to the Nasometer's limited measurements of hypernasality relative to the information that may be used by the rater. They pointed out that the nasalance is derived from the oral-nasal intensity difference at 500 Hz, however the acoustic effects of nasalisation are not restricted to 500 Hz. Watterson et al. (1998b) reported that the Nasometer's best response is between 360 Hz and 600 hz. Spectral analysis of a speaker with nasal emission and hypernasality indicates that there is acoustic energy evident up to 5 K (Appendix 24, first graph). This energy is not analysed in nasometry; however, it is perceived by the rater and may well influence the perceptual judgment of hypernasality.

Voice problems and articulatory errors have been reported to influence perception of nasality (Le Blanc & Shprintzen, 1996; Kuehn, 1982). Changes in fundamental frequency have been found in speakers with velopharyngeal dysfunction (Zajac & Linville, 1989; Sapienza et al., 1996). Watterson et al. (1993) stated that lack of agreements between the Nasometer and perceptual ratings may be partly due to listener's perceptions of nasality being influenced by suprasegmental features and articulatory errors. It is also possible that changes in the acoustic characteristics of voice production may influence nasalance scores if these changes are evident around 500 Hz. However, of the twelve participants identified as having a weak relationship between measurements, only two had articulatory errors and one had mild dysphonia. In the present study voice and articulatory errors did not appear to influence the agreement between perceptual and nasometric results.

If there is a discrepancy between perceptual ratings of nasality and nasometry measurements, then it will be important to evaluate all the influencing factors (consistency of hypernasality, associated nasal airflow errors, voice problems and articulatory errors). If the nasality is inconsistent, the speaker may have a problem with coordination of velopharyngeal closure, in which case palatal surgery may be inappropriate. Further investigations such as nasendoscopy and videofluoroscopy would be indicated and therapy using biofeedback may be appropriate. The presence of associated nasal airflow errors needs to be investigated and its influence on perceptual and nasometry measurements needs to be addressed. Associated voice problems may be highlighted by differences in perceptual and nasometric measurements. This may require detailed voice assessment and therapy to reduce such factors as vocal cord hyperfunction.

In the data collection, silent intervals were recorded in order to identify each sentence category. During analyses of the Total Test sentences there was up to 15 seconds of silence included. The Nasometer Manual (1994) reports that "silence does not affect computed mean nasalance scores" (pg. 102). In order to eliminate the silent intervals as a factor which influences the relationship between measurements, nasometry recordings of ten participants were selected at random. Nasalance scores for the five second gap (i.e., silent intervals) between each sentence category were calculated. Of the thirty silent intervals evaluated, two had nasalance scores above 0%. These nasalance scores were as high as 65% and 70%. This is difficult to explain. Although there was no indication of acoustic energy on the graph, it is possible that breathing or some other noise was detected by the Nasometer and included in the analyses. This would have influenced the relationship between perceptual ratings of nasality and nasalance scores for the Total Test sentences.

The high incidence of HP consonants in the Total speech sample and in the HP consonant sample needs to be considered in analyses. The Nasometer measures oral and nasal acoustic energy in frequencies around 500 Hz. It was therefore anticipated that there should be a nasalance score of zero for HP consonants in speakers without nasal airflow errors. Normative data confirms that there is no significant difference between nasalance scores for HP consonant sentences and LP consonant sentences (Section 3. 3.

2). However, if silent intervals can produce a high nasalance score as above, it is possible that there was also a nasalance score for HP consonants, which had not been anticipated. The effect of the presence of silent intervals and the high incidence of HP consonant sounds in nasometry recordings needs further investigation.

There was a wide scatter of nasalance scores for the participants who were perceived to have mild to moderate hypernasality. This group represents participants with borderline problems (i.e., borderline between perceptually acceptable and unacceptable nasality). Watterson et al. (1996) also found that most of the participants involved in disagreements between perceptual scores and nasalance scores either had borderline perceptual ratings of nasality or borderline nasalance scores. 'Because borderline cases are difficult to classify, absolute nasalance cut-off scores may never be a reality' (Watterson et al., 1996; p. 72). As a result of this disagreement, clinicians must evaluate nasalance scores for the borderline patients with caution. As Watterson et al. (1996) point out, the clinician will have least confidence in the nasalance scores for these patients.

7. 7 Efficiency of Speech Stimuli in the identification of Nasality and Nasal Airflow Errors on the Nasometer

There are considerable benefits in using the Total Test sentences for nasometric assessment in the clinical situation. There is a good relationship between speech assessments on sentence repetition and spontaneous speech (Van Demark, 1964). Hence, sentence repetition is a useful and economic method of data collection (Sell & Grunwell, 2000). The sentences are meaningful and easy to repeat, whilst containing target sounds which are recognised as vulnerable to the effects of velopharyngeal dysfunction (Sell & Grunwell, 2000). Secondly, the sentences are also used as part of a Perceptual Profile, thus allowing for assessment of speech using the same speech stimuli. Thirdly, the Total Test sentences were easily divided into smaller sentence categories to allow for separate evaluation of nasalance scores on HP consonant, LP consonant and Nasal sentences.

Karnell (1995) hypothesised that patients with nasal turbulence would have increased nasalance score on HP consonant sentences compared to their nasalance scores on LP consonant sentences. This hypothesis was supported during data collection in the present study. When recording the speech samples on the Nasometer, sudden increases in nasalance were occasionally detected on the screen when participants produced HP consonants with nasal turbulence. The following graphs indicate the difference in nasalance scores during the production of HP consonants and LP consonants for a participant with nasal turbulence. During the sentence “Gary’s got a bag of lego”, nasal turbulence was perceived on initial /g/ in “Gary’s”. This participant had a nasalance peak of approximately 90% on the initial /g/. On the low pressure consonant sentence “Will you wear a lily” no nasal turbulence was perceived, and the highest peak was approximately 55%, (Figure 7.1).

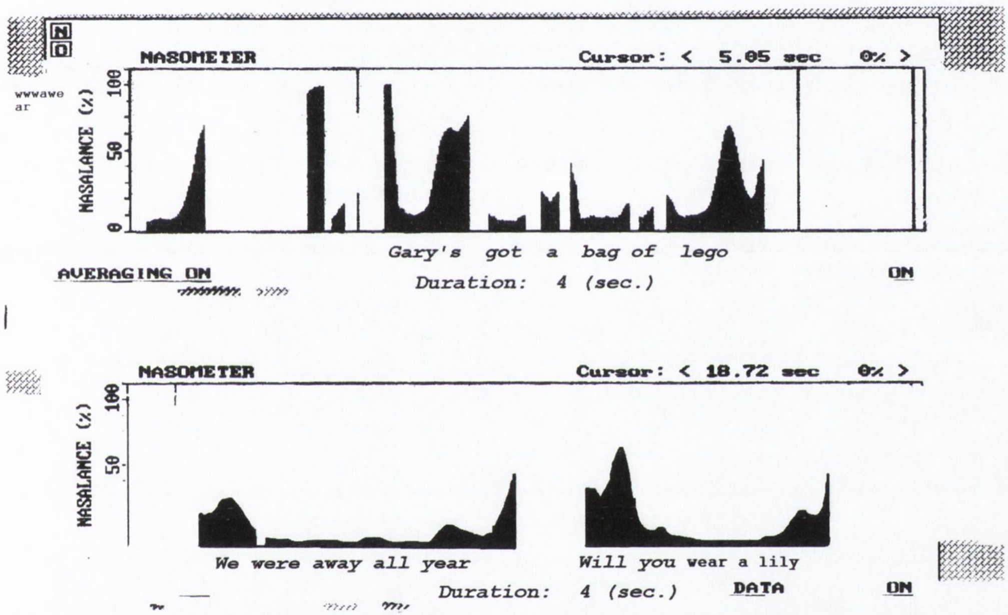


Figure 7.1 The top nasometry graph indicates a high nasalance peak of 90% for a participant with nasal turbulence during the production of a HP consonant /g/. The bottom graph indicates the highest nasalance peak of 55% during the production of low pressure consonants for the same speaker.

The difference in nasalance peaks indicated that there may be an increase in nasalance scores on HP consonant sentences compared to LP consonant sentences for participants with nasal turbulence. Spectrographic analyses of HP consonant and LP consonant sentences of speakers with nasal turbulence suggest that there was low frequency acoustic energy associated with nasal turbulence. Recorded data from each of the nasal airflow categories in the present study were randomly selected for spectral analyses. Spectral analyses indicated that four speakers (two with perceptible nasal turbulence and two with a velopharyngeal fricative) had low frequency energy at approximately 100 to 300 Hz on the spectrograms (Appendix 24). This low frequency energy was not evident in another four speakers (two with nasal emission and two with nasal fricatives) when no 'snorting' sound was present. If the Nasometer detected this low frequency energy associated with nasal turbulence, the nasalance scores may have been elevated for these speakers.

To evaluate the influence of nasal turbulence and/or nasal emission on nasalance scores, participants with nasal turbulence/nasal emission were identified and the mean nasalance scores for the HP consonant sentences and the mean nasalance scores for LP consonant sentences were calculated (Table 7. 2). The difference between the mean nasalance scores for the two sentence categories was significant. But, contrary to Karnell's hypothesis, nasalance scores on HP consonant sentences were lower than nasalance scores for LP consonant sentences for participants who had perceived nasal emission and/or nasal turbulence.

Table 7.2 Mean nasalance scores for participants with perceived hypernasality and nasal turbulence, and for participants with perceived hypernasality and nasal emission during production of HP consonant sentences and LP consonant sentences. Significant differences between scores for HP consonant and LP consonant sentences are found.

Participants	**Mean Nasalance HP**	**Mean Nasalance LP**	**t score**	**p =**	**df**	**95% CI**
Hypernasal + Nasal Turbulence group	35%	41%	-2.83	.013	16	-8.9, -2.4
Hypernasal + Nasal Emission group	39%	45%	-3.62	.001	26	-8.9, -2.4

Karnell (1995) divided his population into three groups using a cut-off value of 31%.

- Group 1 had HP consonant and LP consonant nasalance scores above the cut-off value and for this group there was no significant difference between HP consonant and LP consonant scores.

- Group 2 had HP consonant and LP consonant nasalance scores below the cut-off value. There was a small but statistically significant difference between HP consonant and LP consonant scores.
- Group 3 had mixed results.

Karnell (1995) reported that five out of ten participants in group 3 in his study, had nasalance scores above the cut-off value for HP consonant sentences but not for LP consonant sentences. Karnell hypothesised that this group had nasal turbulence on HP consonants, thus elevating the nasalance score. Karnell did not assess this group perceptually to confirm this hypothesis.

In order to compare results of the present study with Karnell's study, a cut-off value of 24% for HP consonant sentences and 28% for LP consonant sentences was identified. This was calculated using the statistically determined cut-off value (i.e. 2 standard deviations above the mean for normal speakers for the production of HP consonant and LP consonant sentences). Similar groupings were made to Karnell's study.

- Group 1 (31 out of 50 participants) had nasalance scores above the cut-off for both HP consonant and LP consonant sentences. The difference between the mean nasalance scores was statistically significant ($t = 3.52$, $p = .002$).
- Group 2 (14 out of 50 participants) had nasalance scores below the cut-off for both HP consonant and LP consonant sentences. The difference was not significant ($t = .84$, $p = .417$), but as in Karnell's study, the difference was small.
- Group 3 (5 out of 50 participants) had mixed results. Four participants had LP consonant scores above the threshold and HP consonant scores below the threshold and one participant had HP consonant scores above the threshold and LP consonant scores below the threshold.

In the above groupings, all the results compared well with Karnell's results. However, only one participant in the present study had HP consonant scores above the cut-off and LP consonant scores below the cut-off. This is the group that Karnell hypothesised had nasal turbulence which may have elevated the nasometric scores. But, seventeen participants had perceived nasal turbulence in the present study and only one was in this group. In the present study, calculating separate nasalance scores for HP and LP consonant sentences did not identify participants with nasal turbulence. Watterson et al.

(1998a) assessed a group of participants with varying degrees of nasality and visible nasal emission (detected by the mirror test). They reported marginally higher nasalance scores on HP consonant sentences compared to LP consonant sentences, but the difference was not significant. One drawback of the Watterson et al. (1998a) study is that no perceptual assessment of nasal airflow was undertaken. Results of the present study, which included perceptual assessment of nasal emission/turbulence, confirm Watterson et al.'s finding that no clinical or diagnostic information was gained by calculating separate nasalance scores for HP and LP consonant sentences.

There are two possible reasons for the lack of significant differences in the nasalance scores for HP consonant sentences and LP consonant sentences for speakers with nasal turbulence. One is that the Nasometer may not have detected the low frequency energy that was possibly associated with nasal turbulence (at around 100 Hz to 300 Hz), as its best response is above 360 Hz, and it may have filtered out these lower frequencies. Although, higher nasalance peaks were observed on HP sounds in speakers with nasal turbulence during data collection, these increases were not seen consistently. Further investigation into the acoustic correlates of nasal turbulence is recommended.

Another explanation may be that there was increased articulatory effort during the production of HP consonants. It is possible that participants may have benefited from an increased level of velopharyngeal closure usually associated with pressure consonant production (Karnell, 1995). This increased velopharyngeal closure may have resulted in better velopharyngeal closure during the production of vowels adjacent to pressure consonants. Hence, there was improved velopharyngeal function during the production of the HP consonant sentences compared to the LP consonant sentences.

The nasalance measurement during the production of the nasal sentence was found to be a valid indicator of perceived hyponasality. However, reliability of the Nasal sentence was poor. This was probably due to the limited number of nasal sentences used in the assessment. The weak reliability results support the findings of Nichols (1999) that speech stimuli with less than 5 sentences are unreliable. Hence, for clinical use a more detailed Nasal stimulus will be required.

7. 8 Summary of Nasometry Results

Results of the present study indicated that the comprehensive and phonetically balanced speech sample used with the Nasometer distinguished between normal and pathological speakers. The Total Test sentences and the HP consonant sentences were good indicators of hypernasality, while the Nasal sentence was a good indicator of hyponasality. However, there was no benefit in obtaining separate nasalance scores for LP consonant sentences. Analyses of cut-off values between normal and abnormal nasality indicated that a range of cut-off values had the greatest efficiency in identifying normal and pathological speakers, in a similar manner to perceptual judgements. Stronger relationships between perceptual ratings of nasality and nasalance scores found in this study are probably due to the descriptive nature of the perceptual scale and its good intra-rater reliability. Factors which influence the relationship between perceptual and nasometric measurements will have clinical implications. If there are discrepancies between measurements, then factors such as consistency of hypernasality and associated nasal airflow errors need to be considered. Further investigations into acoustic characteristics of nasal airflow errors and the effects of articulatory effort on velopharyngeal function may provide valuable information.

7. 9 Pressure/Flow Measurements

The first section (7.10) reviews the pressure/flow results obtained in this study and compares them to normal results. The normal scores were obtained from 152 normal speaking children in America (Zajac, 1998, personal communication). Results of the nasal flow measurements, oral pressure, differential pressure, and velopharyngeal area measurements from the study population were compared with normal scores. Results are discussed for each measurement. The relationship between pressure/flow measurements and perceptual assessment of nasal airflow problems are discussed in the following section (7.11). The final sections discuss the cut-off values which can be used to distinguish between normal and abnormal speakers (7.12), the factors which may influence the relationship between measurements (7.13) and the value of the speech samples used in the pressure/flow assessment (7.14).

7. 10 Effectiveness of Pressure/Flow Measurements in the Discrimination of Normal and Pathological Speakers

The descriptive statistics indicated a skewed distribution of the pressure/flow measurements. As a result, the mean measurements for the study group were somewhat inflated. However, due to the large sample size it was possible to compare differences in mean pressure/flow results from the study group with the means from normal speakers using independent sample t-tests. Normal pressure/flow measurements were provided by Zajac et al. (1998, personal communication), but analysis of variance could not be carried out, as only normal means and standard deviation scores were available.

7. 10. 1 Nasal Flow

Results indicated that nasal flow measurements during the production of /p/ in ‘pa’, ‘pi’ and ‘hamper’ were able to distinguish between a group of normal speakers and a group with suspected velopharyngeal dysfunction. For all three speech stimuli, there was greater mean nasal airflow and a greater standard deviation score for the study group compared to normal speakers. Results indicated an extremely wide range of nasal flow measurements for the study group, with some measurements as high as 300 to 400 mls.

7. 10. 2 Oral Pressure

Mean oral pressure measurements for the study group, during the production of the three speech stimuli, were marginally lower than the mean oral pressure measurements for normal speakers. Although the differences between groups of speakers were small, they were significant for the production of /p/ in 'pa' and /p/ in 'hamper'. However, the mean oral pressure measurement for the study group was within the normal range. Interestingly, 4 participants had extremely high oral pressure measurements (10 to 14 cm H_O), despite the presence of hypernasality. These results support previous studies which found oral pressure measurements greater than 3 cm H_2O in speakers with velopharyngeal incompetency (Dalston et al., 1988; Warren, 1989). Dalston et al. (1988) suggested that pressure maintenance resulted from changes within the vocal tract. This maintenance of oral pressure, according to Warren (1986), is a primary goal of the vocal tract system; however, why this happens is not understood.

The standard deviation scores during the production of all speech stimuli were similar for normal and pathological speakers. However, the range of oral pressure measurements was greater for the pathological speakers. One could not be confident in using oral pressure measurements to distinguish between normal and abnormal speakers.

7. 10. 3 Differential Pressures

None of the aerodynamic studies of normal children or adults obtained normal differential pressure measurements. Hence, comparison with normative data was not possible. Previous studies have indicated that differential pressures below 3 cm H_2O were indicative of velopharyngeal dysfunction. Results from the present study indicated that mean differential pressure measurements for the study group were greater than 3 cm H_2O for all 3 speech stimuli. Differential pressure measurements ranged from 0.03 to 19 cm H_2O, indicating that some participants had differential pressure measurements below the norm.

7. 10. 4 Velopharyngeal Port Area

Results from the present study indicated significant differences between mean velopharyngeal area measurements during the production of /p/ for normal speakers and speakers with varying degrees of velopharyngeal dysfunction. This indicates that velopharyngeal port area measurements can distinguish between normal and pathological speakers. Standard deviation measurements also indicated a marked difference between normal and abnormal speakers. An extremely wide range of velopharyngeal port area measurements was found on all three speech stimuli, with 6 participants having extreme values. These findings indicate a skewed distribution of velopharyngeal area measurements, with a large difference between mean and median measurements (e.g. mean for /p/ in 'pa' = .083 and median of /p/ in 'pa' = .008).

In summary, the present findings indicated that nasal flow measurements and velopharyngeal port area measurements for /p/ distinguished between normal and pathological speakers. Oral pressure measurements were poor discriminators of normal and pathological speech. The discriminative value of the differential pressure measurements requires further analysis, along with the acquisition of normative data.

7. 11 Relationship between Pressure/Flow Measurements and Perceptual Ratings of Nasal Airflow

In order to analyse the results of the perceptual assessments of nasal airflow errors, a numerical scale was designed as follows:

weak, inconsistent, infrequent	=	1
weak, inconsistent, frequent	=	2
weak, consistent, infrequent	=	3
weak, consistent, frequent	=	4
strong, inconsistent, infrequent	=	5
strong, inconsistent, frequent	=	6
strong, consistent, infrequent	=	7
strong, consistent, frequent	=	8

This allowed for all possible descriptions of nasal airflow errors. In this way, the descriptive scale was ranked from 1, indicating mild errors, to 8, indicating severe errors. However, the scale was not equidistant or ordinal. The difference between weak consistent and frequent errors was not necessarily one point below strong, inconsistent and infrequent errors. The difference between consistent and infrequent errors was marginally different from inconsistent and frequent errors. In order to evaluate the ranking of the scale, each description was weighted, so that the strength of airflow was given double points (e.g. weak was assigned 2 points and strong was assigned 4 points, while inconsistent and infrequent were assigned 1 point, and consistent and frequent were assigned 2 points). Correlation coefficients were calculated for the relationship between the unweighted scales and perceptual ratings of nasal airflow errors. Further correlation coefficients were calculated for the relationship between the weighted scales and the perceptual ratings of nasal airflow errors. Results indicated no difference in correlation coefficients between perceptual ratings of nasal airflow and pressure/flow measurements, using the weighted and unweighted scales. This indicated that the ranking scale used above was adequate for the statistical analysis of the data.

The relationship between pressure/flow measurements and perceptual assessment of nasal airflow is discussed under each category of nasal airflow errors. Correlational analysis is discussed and, where indicated, further assessment of the relationships is evaluated using test sensitivity and specificity. Discussion of the appropriate cut-off values for children is presented. None of the studies in the literature examined the relationship between pressure/flow measurements and perceptual judgements of nasal airflow.

7. 11. 1 The Relationship between Pressure/Flow Measures and Perceptual Ratings of Nasal Emission

Correlational analysis using Spearman Rank was used to ascertain if the two assessment techniques rated participants in similar manner. Correlation coefficients between the nasal emission and pressure/flow measurements indicated varying results. Sensitivity and specificity used statistically determined cut-off values and then calculated overall efficiency. Clinically determined cut-off values are discussed below.

7. 11. 1 (i) The Relationship between Perceptual Ratings of Nasal Emission and Nasal Flow Measures

A moderate positive significant relationship was found between nasal flow and nasal emission during the production of /p/ in the word hamper. The relationship for /p/ in 'hamper' was greater than the relationship for /p/ in 'pa' and 'pi'. This was expected, as participants with nasal emission would be expected to exhibit this problem during the production of a nasal/plosive cluster. This cluster requires the rapid velopharyngeal closure, which is similar to the closure required for fricative and plosive production in continuous speech (Warren, 1979).

Results indicated a high test sensitivity of 0.92 for nasal flow measurements on /p/ in 'pa', using a cut-off value of 4 ml/s. This cut-off was statistically determined based on normal mean nasal flow of 1.2 ml/s and a standard deviation of 1.4 ml/s (Zajac 1998, personal communication). A weak test specificity of 0.43 was found. This indicated an overall efficiency of 0.69. The poor test specificity may be explained by the presence of inaudible nasal emission. Inaudible nasal emission exists where there is evidence of nasal airflow on a mirror or some instrumental measure but the emission is not perceptible (Bzoch, 1989; Sell et al., 1994). The study group consisted of participants with a history of cleft palate and/or velopharyngeal dysfunction. It is likely that many participants had inaudible nasal emission. Therefore, they may have had nasal flow measures above 4 ml/s, but this was not perceived in the perceptual assessment.

During the production of /p/ in 'hamper' results indicated excellent test sensitivity (0.96). Specificity was 0.62 and overall efficiency was 0.79. This analysis indicated that the speech sample 'hamper' had better sensitivity, specificity and overall efficiency than the speech samples 'pa' and 'pi'.

The low specificity of nasal flow measurements, which was probably due to the presence of inaudible nasal emission, has clinical value. If a child presents with no perceptible nasal emission, but has increased nasal flow measurements, he/she may require regular monitoring of speech, as he/she may be at risk for development of nasality and/or nasal airflow problems with growth, adenoidal involution or surgery for maxillary advancement.

These findings suggest that the nasal flow measurement using the PERCI SARS system is a good indicator of perceptually determined nasal emission. The results indicate that nasal flow measurements on /p/ in hamper are the most efficient indicator of nasal emission.

7. 11. 1 (ii) The Relationship between Perceptual Ratings of Nasal Emission and Velopharyngeal Port Area

A good significant correlation was found between nasal emission and velopharyngeal port area during the production of /p/ in the word 'hamper'. Moderate significant relationships were found between velopharyngeal port area and nasal emission during the production of the other speech stimuli.

Results indicated excellent agreement between the presence of nasal emission and velopharyngeal port area (sensitivity = 0.92) during the production of 'pa'. Again specificity was low (0.43), but overall efficiency was acceptable. Results improved for the speech stimulus 'hamper' (sensitivity = 0.96), with definite improvement in test specificity (specificity = 0.59) and overall efficiency (0.79). Results indicated that velopharyngeal port area measurements, during the production of the word 'hamper', can discriminate between the presence and absence of nasal emission.

7. 11. 1 (iii) The Relationship between Perceptual Ratings of Nasal Emission and Pressure Measurements

Weak correlations were found between the nasal emission and oral pressure measurements for all speech stimuli. The relationship between differential pressure and nasal emission was moderate to good (ranging from -0.60 to - 0.68). Interestingly, differential pressure measurements indicated a negative relationship. One would expect pressure measurements to have a negative relationship with nasal emission, because as nasal emission increases one would expect oral pressure or differential pressures to decrease.

Weak test sensitivity of 0.19, test specificity of 1.0 and overall efficiency of 0.57 were found when comparing oral pressure measurements on 'pa' and perceptual assessment of nasal emission. Sensitivity and specificity results during production of /mp/ in hamper were marginally better, with a sensitivity of 0.30, a specificity of 1 and an overall efficiency of 0.62. These results suggest that oral pressure measurements were not a good indicator of the presence of nasal emission.

A moderate relationship was found between differential pressure measurements and nasal emission. During production of 'pa', test sensitivity was 0.44, specificity was 0.95 and overall efficiency was 0.66. The weak test sensitivity indicated that differential pressure measurements during the production of 'pa' did not identify participants with nasal emission. A stronger relationship was found between perceptual and differential pressure measurements during the production of /p/ in hamper (test sensitivity of 0.72, specificity of 0.81 and overall efficiency of 0.76). These findings suggest that differential pressure measurements during the production of /p/ in 'hamper' are a good indicator of nasal emission.

7. 11. 2 The Relationship between Pressure/Flow Measures and Perceptual Rating of Nasal Turbulence

A correlation analysis indicated weak negative relationships between nasal flow, velopharyngeal port area and perceptual assessment of nasal turbulence. There was a positive significant relationship between oral pressure measurements and nasal turbulence. This was unexpected, as one would expect oral pressure to decrease as nasal turbulence associated with velopharyngeal dysfunction increased. Differential pressures had a positive non-significant relationship with nasal turbulence. These findings suggest that pressure/flow measures are a poor indicator of nasal turbulence. Because of the weak relationship between nasal turbulence and pressure/flow measurements, no further analysis was carried out.

The weak relationship between pressure/flow measurements and perceptual ratings of nasal turbulence may support some theories regarding the articulatory production of nasal turbulence. Trost (1981) reported velar activity during the production of nasal turbulence. Riski (1999, personal communication) stated that during nasal turbulence there was light contact between the velum and pharyngeal walls resulting in a velar trill. If nasal turbulence is a velar trill, then there may be sufficient closure of the velopharyngeal sphincter to maintain oral pressure and prevent significant nasal flow. Hence one would expect adequate oral pressure, no significant increase in nasal flow and velopharyngeal port area, as was found in the present study.

Another explanation for the weak relationship between measurements may be the small number of participants in the study group who presented with perceived nasal turbulence. Thirty four percent (n = 17) of the group had perceived nasal turbulence, all of which presented with weak nasal turbulence. It is possible that mild errors were not detected during the pressure/flow assessment. Of the 17 cases with nasal turbulence, 8 had inconsistent nasal turbulence. If the errors were inconsistent, they may not have been evident during the limited speech sample of the pressure/flow assessment, but may have been evident during the perceptual analysis.

The limited speech sample used in pressure/flow measurements may result in the inability of the instrument to detect nasal turbulence. Weak nasal turbulence is more likely to be detected during conversational speech. Since nasal turbulence is usually associated with a very small velopharyngeal gap (Kummer et al., 1992), this gap may not be evident during the single word repetitions. McWilliams and Philips (1979) stated that the production of consonants in small articulatory units (i.e. syllables or isolated words) demands less of the velopharyngeal valving mechanism than production of consonants in more difficult articulatory contexts. They reported a case where the child could achieve adequate velopharyngeal closure during production of a single word but not during the production of rapid spontaneous speech. As the participants in this study had mild nasal turbulence (i.e. weak), it is possible that they could achieve adequate velopharyngeal closure during the production of the limited speech sample. However, the nasal turbulence may have been perceived during sentence repetitions and conversational speech.

Nasal turbulence is usually associated with voiceless plosives and fricatives (Sell et al., 1998). The limited speech sample did not allow for assessment of fricative sounds. It is also possible that nasal turbulence was detected only on fricative sounds, which were not assessed during pressure/flow measurements. Zajac et al. (1996) attempted to assess the fricative sound /s /using the pressure/flow technique. They reported that results for this assessment were unreliable due to problems of placement of the pressure catheter in the oral cavity.

The perceptual assessment of nasal turbulence was found to have poor intra and inter-rater reliability using the Perceptual Profile. This was explained by the high incidence of weak nasal turbulence in the reliability study population. In the present study, all participants also had weak nasal turbulence. Hence, it is possible that the reliability of the perceptual rating of nasal turbulence was questionable. This may have influenced the relationship between the pressure/flow measurements and perceptual measurements.

7. 11. 3 The Relationship between Pressure/Flow Measures and Perceptual Ratings of Nasal Fricatives

Correlation analysis of pressure/flow measures and perceptual ratings of nasal fricative indicated extremely weak, non-significant relationships between measurements. This may well be explained by the fact that only eight participants in the study had a nasal fricative. In seven cases the nasal fricative was phoneme specific on targets /s/ and /z/ and, therefore not present during the production of the speech sample used in pressure/flow measurements. As a result of the limited speech sample, one would not expect a good correlation between perception of nasal fricatives and pressure/flow measures. Further analysis was not indicated because of the weak relationship.

7. 11. 4 The Relationship between Pressure/Flow Measures and Perceptual Ratings of Velopharyngeal Fricatives

Results of correlation studies between pressure/flow measurements and perceptual ratings of velopharyngeal fricatives indicated weak relationships between the two assessment techniques. Interestingly, a negative but significant correlation was found between nasal flow measures and velopharyngeal fricative and between velopharyngeal port area and velopharyngeal fricative. The fact that the relationship was negative was unexpected, as one would have expected an increase in nasal flow and small gap in the velopharyngeal port with the presence of a velopharyngeal fricative. This finding is difficult to explain. More importantly, 6 out of the 7 participants had a phoneme specific velopharyngeal fricative which was not present on /p/ and therefore not detected by pressure/flow measurements.

The limited speech sample used for pressure/flow measurements is a significant problem in the instrumental assessment of nasal airflow errors associated with fricatives and plosives.

7. 12 Efficiency of Cut-off Values

Only cut-off values used in the assessment of test sensitivity and specificity for the detection of nasal emission were evaluated. Other nasal airflow problems were not analysed, due to the weak relationships found in the correlational analysis.

In order to ascertain the optimal cut-off value for nasal emission and pressure flow results, sensitivity and specificity were calculated using different cut-off values. The value that produced the optimal overall efficiency value was identified.

Analysis of sensitivity and specificity for nasal flow measures during the production of /pa/ for the identification of nasal emission indicated the following:

- Using a cut-off value of 9 ml/s, which represents the maximum range for normal speakers, sensitivity was lower (0.80), and specificity increased to (0.56). However, overall efficiency remained at 0.69;
- A greater cut-off value of 20 ml/s produced a sensitivity score of 0.73, a specificity score of 0.78 and overall efficiency score of 0.75.

This would indicate that the clinically determined cut-off value of 20 ml/s is most efficient in identifying the presence or absence of audible nasal emission. However, the statistically determined value of 4 ml/s produced excellent sensitivity of 0.92. As nasal emission can be present but inaudible, the specificity and resulting overall efficiency may be reduced. Using a range of cut-off values (4 to 9 ml/s) for this speech sample would be most beneficial in assessing nasal emission.

For nasal flow measures during the production of /p/ in hamper:

- A cut-off value representing the maximum range of normal scores (39 ml/s) resulted in a decreased sensitivity of 0.73, an increased specificity of 0.85 and the same overall efficiency of 0.78;
- A cut-off value of 30 ml/s produced a sensitivity score of 0.88, the specificity of 0.80 and an improved overall efficiency of 0.85.

Therefore, a cut-off range between 14.5 ml/s and 30 ml/s during the production of /p/ in hamper would be most efficient in identifying the presence or absence of nasal emission.

In the analysis of velopharyngeal port area during the production of /pa/ and perceptual judgements of nasal emission, varying cut-off values resulted in increased overall efficiency.

- Using a cut-off of value of .004 cm^2 sensitivity decreased to 0.88, and specificity increased to 0.60 and overall efficiency was 0.75.
- The greatest overall efficiency was found using cut-off value of 0.01 cm^2. This produced a sensitivity of a 0.73, a specificity of 0.86 and an overall efficiency of 0.79.

These results would indicate that a range of cut-off values from .004 to .01 cm^2 should adequately discriminate between the presence and absence of nasal emission. There was no change in overall efficiency scores with the change in cut-off values for velopharyngeal port area during the production of /p/ in 'hamper'.

Because of poor test sensitivity for differential pressure measurements during the production of 'pa', no evaluation of optimum cut-off values was undertaken. In order to ascertain the optimum cut-off value for differential pressures during the production of /p/ in 'hamper', varying cut-off values were used. As no normative scores for differential pressure measurements were available, varying cut-off values above and below 3 cm H_2O were evaluated. Using a cut-off value of 4 cm H_2O, a sensitivity of 0.92, a specificity of 0.68 and an overall efficiency of 0.80 was found. Hence, a differential pressure on /mp/ between 3 and 4 cm H_2O, may indicate the presence of nasal emission.

7. 13 Factors which Influence the Relationship between Pressure/Flow Measures and Perceptual Ratings of Nasal Airflow Problems

Test-retest analysis of pressure/flow measurements indicated substantial individual variation in repeated measurements of nasal flow and velopharyngeal port area for speakers (see Table 6.10 & 6.11). Greater variability of individual measurements was found for pathological speakers compared to normal speakers. Certain factors in the methodology of this study may have contributed to this individual variability in scores. Firstly, the elicitation procedure involved nine repetitions of the /p/ sound. The aim was to base measurements on six productions of /p/, so that the mean measurement could be calculated. This was in order to overcome some of the normal individual variations in sound production. It has been recommended that the initial /p/ in the speech sample should not be measured, as this sound production is highly variable (Warren, 1995, personal communication). However, many participants did not repeat nine syllable/words continuously and required prompting from the examiner after three repetitions. This may have given rise to listing effects where each syllable was produced with different degrees of articulatory and/or respiratory effort, resulting in wide variation of pressure and flow during production of the syllables. A second factor in the methodology was that there was no stringent control for variation in loudness during syllable/word production. If the examiner noted that the participant was using increased loudness during recording, the participant was asked to repeat the speech sample in a quieter voice. This control for loudness was subjective and may not have been reliable. Variations in loudness may have increased individual variability in measurements. Finally, participants may have made greater articulatory effort during production of some syllables and not others. No attempt was made to control for this. The larger variation in measurements for pathological speakers compared to normal speakers may be explained by increased articulatory and respiratory effort. Whereas participants with velopharyngeal dysfunction may be able to produce isolated sounds normally, they may be unable to maintain velopharyngeal function of a string of utterances (McWilliams & Philips, 1979). In this study, participants may have used increased articulatory effort during some syllable/word productions in the pressure/flow assessment, in an unconscious effort to overcome this.

Inconsistency of sound production, however, may be a function of the nasal airflow speech errors. Perceptual analyses indicated that almost half (48%) of the nasal airflow errors perceived were inconsistent. In order to evaluate this as a factor for assessment of nasal emission, Spearman Rank correlations were calculated for participants with consistent errors only. Therefore when participants with inconsistent nasal airflow were omitted from the calculation, the correlation coefficient increased from 0.5 to 0.7. This would indicate that consistency of nasal airflow errors influenced the relationship between perceptual ratings of nasal emission and pressure/flow measurements.

Analysis of the pressure/flow graphs indicated that 17 (35%) of the participants had simultaneous nasal pressure, nasal flow and oral pressure peaks during the production of the word 'hamper' (Figure 7.2).

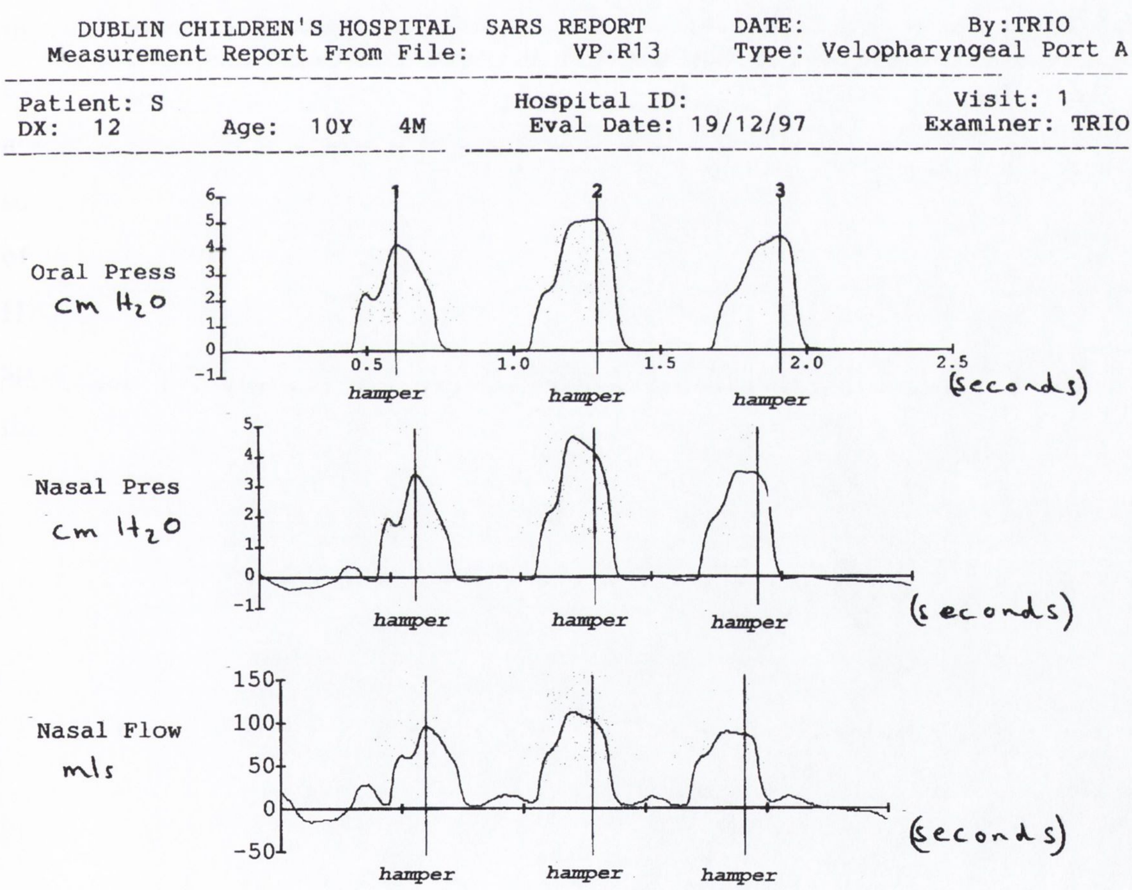
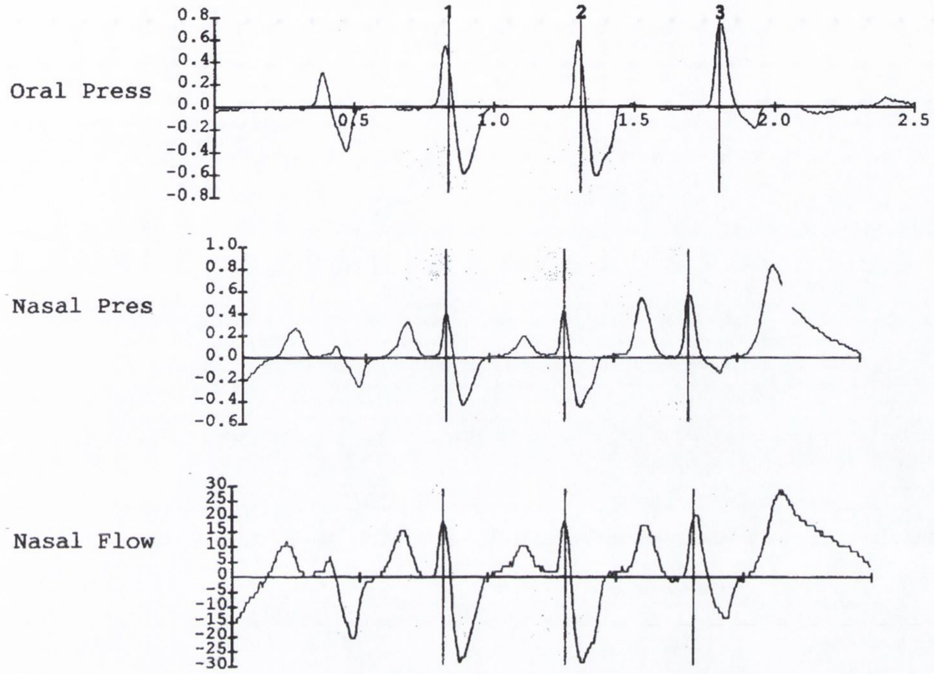


Figure 7.2 PERCI SARS graph indicating oral pressure, nasal pressure and nasal flow graphs, during three productions of the word 'hamper'. This graph indicates an overlap of the peak oral pressure and peak nasal flow and peak nasal pressure.

In normal speech production of 'mp' the peak nasal flow is before the peak oral pressure indicating the /m/ sound prior to the release of the /p/ in 'hamper' (see Appendix 19c). Possible explanations for this overlap in pressure/flow tracing were considered. Segmentation of the cluster /mp/ in the bisyllabic word may have varied considerably between speakers which may have influencedthe way the consonant cluster was produced. For example, Abercrombie (1967) stated that there is considerable variation in syllable division in normal speakers. Little is known about the effects of variation in segmentation on pressure/flow measurements and indeed their effects in a pathological population. It was hypothesized that participants with velopharyngeal dysfunction did not signal the oral /p/ in the /mp/ cluster. Therefore the graphs and audio signals of the 17 participants with simultaneous pressure and flow peaks were re-evaluated in order to perceptually assess the segmentation of the /mp/ cluster. Analysis indicated that all of these participants did signal the /m/ and the /p/ in the cluster. The oral peak pressure evident on each graph, albeit a low peak pressure in some cases confirmed this. The auditory signal indicated that 4 out of the 16 participants had weak production of the /p/ sound. One graph indicated that the participant had a negative trace following the release of the /p/ sound in 'hamper' and had low oral pressure measurements of 0.5 to 0.7 cm H_2O (Figure 7.3). Reassessment of the graphs and the audio signal indicated that this speaker had glottal coarticulation during the production of /p/ and this may have caused the weak oral pressure and the unusual negative readings on all three channels.

```
DUBLIN CHILDREN'S HOSPITAL  SARS REPORT      DATE:             By:TRIO
Measurement Report From File: C     VP.R04   Type: Velopharyngeal Port A
--------------------------------------------------------------------------
Patient:                              Hospital ID:               Visit:
DX:  0        Age:  0Y    0M         Eval Date: 07/11/97      Examiner: TRIO
--------------------------------------------------------------------------
```



Cur#			VP Area			
1	OP=00.5	NP=00.4	NFlow=018.7	DP=00.14	AREA= 0.054	X=1.66
2	OP=00.6	NP=00.4	NFlow=018.7	DP=00.16	AREA= 0.051	X=2.14
3	OP=00.7	NP=00.6	NFlow=018.7	DP=00.18	AREA= 0.049	X=2.64

VP Area ==================

Oral Press	Mean = 0.6	STD = 00.10	SE= 00.06
Nasal Pres	Mean = 0.5	STD = 00.09	SE= 00.05
Nasal Flow	Mean = 18.7	STD = 00.00	SE= 00.00
Computed DP	Mean = 0.16	STD = 00.02	SE= 00.01
AREA(sq cm)	Mean = 0.052	STD = 0.003	SE= 0.002

Figure 7.3 Pressure/flow graph and measurements for one participant with glottal coarticulation during the production of the word ‘hamper’. Note the negative dip of the graphs on all three channels. (Cursor #1 for the oral pressure peaks shifts to the right in the printed out. This was occasionally noted when the reports were printed. However the oral peak pressure points were based on the readings in the measurement box on the computer screen (See Methodology, Figure 5.6).

Problems in the timing of velopharyngeal closure may also result in overlap between nasal flow and oral pressure graphs. Warren et al. (1993a) found that speakers with velopharyngeal dysfunction had shorter time gaps between the beginning, peak and end of the nasal flow graphs and the beginning, peak and end of the oral pressure graphs than normal speakers (see Section 2. 7. 9). It is possible that the participants in the present study had timing errors of velopharyngeal closure, which resulted in the simultaneous, peak pressure and flow measurements. Further investigation into timing errors is recommended.

In all pressure/flow assessments, a possible source of error was the appropriate placement of tubes. During assessments it was noted that the nasal flow tube tended to move, especially in cases where the child moved his/her head during assessment. Although the tubes were regularly checked during assessments, on some occasions there may have been minimal leakage of air.

Adaptive changes within the vocal tract system have been reported to take place in order to maintain oral pressure levels above 3 cm H_2O (Warren, 1986; Dalston et al., 1988; Warren, 1989). Such changes include increased respiratory effort, nasal grimace, high lingual carriage and increased glottal resistance. Anomalies such as increased nasal resistance, which is common in children with cleft palate, may also result in structural changes within the vocal tract. Such factors were not measured in the present study and may have had a significant influence on the relationship between pressure/flow measurements and perceptual ratings of nasal airflow errors. These will need to be evaluated in clients in the clinical situation, especially when there is a discrepancy between perceptual and instrumental measurements.

There is evidence from aerodynamic assessments that speakers with voice problems associated with velopharyngeal dysfunction may have changes in transglottal pressure and resistance. Lewis et al. (1993) found that a strained or strangled voice quality was associated with high transglottal pressure and resistance, while breathy voice patterns were associated with low transglottal pressure. Five (10%) of the participants assessed in the present study were found to have dysphonia and/or reduced vocal volume. It is possible that such changes in pressure in the vocal tract affected the pressure and airflow within the oral and nasal cavities, thus influencing measurements.

The degree of speech problems of the population studied may explain the weak relationships between pressure/flow measurements and nasal airflow problems. All participants who had perceived nasal emission were rated as having weak nasal emission, while 40% of participants were rated as inconsistent and 51% were rated as infrequent.

This weak, inconsistent and infrequent nasal emission indicated that these participants had borderline or mild velopharyngeal dysfunction. The large number of borderline cases in the study group may have resulted in weak relationships between pressure/flow and nasal airflow problems (Warren, 1998, personal communication). On analysis of the reliability of Perceptual Profile, the strong-weak parameter had poor reliability for nasal emission and nasal turbulence. This emphasizes the difficulty in assessing weak nasal airflow problems reliably both with instrumentation and on a perceptual basis.

The factors which influence the relationship between perceptual and pressure/flow measurements have significant clinical implications as they may influence the type and outcome of management required. Inaudible nasal emission needs to be identified especially in speakers with repaired cleft palate who will need surgery to advance the maxilla following the completion of facial growth. Such patients are at risk of developing velopharyngeal inadequacy postoperatively (Bradley, 1989). If the nasal airflow errors are inconsistent, speech therapy may be indicated to develop consistent velopharyngeal closure during speech. If a child uses increased articulatory and respiratory effort during speech, this will need to be addressed in order to prevent damage to the vocal cords, and ensure healthy voice production. Discrepancies between perceptual and pressure/flow measurements may indicate that these important factors exist and need to be addressed.

7. 14 Speech Stimuli in Pressure/Flow measurements.

As expected, mean nasal flow measurements and mean velopharyngeal port area measurements were greater for /p/ in the word hamper. This is due to the nasal/plosive cluster /mp/, which requires rapid opening and closing of the velopharyngeal sphincter. The range of nasal flow measurements for /mp/ was less than the range for /pa/. However, only one participant had extreme nasal flow measurements on /pa/. When this participant was eliminated from the group, nasal flow measurements for /mp/ were greater than nasal flow measurements for /pa/.

It had been anticipated that the strongest relationship between perceptual measures of nasal emission and pressure/flow measures would be found for the production of the word 'hamper'. Warren (1979) reported that the /mp/ cluster in the word hamper best represented conversational speech in aerodynamic studies. The present findings support this.

Smith and Guyette (1996) reported that an excessive velopharyngeal opening was more evident on /pi/ than on /pa/. The nasal flow measurements and velopharyngeal port area measurements in the present study do not support this finding. Ten participants had increased nasal flow on /pa/ compared to /pi/, while six participants had increased nasal flow on /pi/ compared to /pa/. Of the group with increased scores on /pa/, seven were perceived to have inconsistent nasal emission and/or nasal turbulence. This inconsistency may explain the variation in score between the two speech stimuli. These results would indicate that both speech stimuli should be included in the test sample.

With respect to velopharyngeal port area measurements, results indicated that the mean velopharyngeal port area and the range of area measurements was greater for 'pa' than for 'pi'. Smith and Guyette (1996) identified eight participants, out of a total of fifty one, who had consistent variability in velopharyngeal opening and closing. All eight participants had openings during 'pi' repetitions but not during 'pa' repetitions. In the present study, twenty participants had variability of velopharyngeal port area measurements for 'pa' and 'pi'. However, of these, only eight of the 20 had greater velopharyngeal port area during the production of 'pi' compared to 'pa'. Analysis of measurements for participants with borderline velopharyngeal closure (i.e. around .005 cm^2) during the production of either sound indicated that only one participant had greater area measurements on 'pi' compared to 'pa'. This does not support Smith and Guyette's (1996) claim that participants with borderline velopharyngeal function have greater velopharyngeal openings during the production of 'pi' compared to the production of 'pa'.

The findings from the three speech samples suggest that the nasal flow and velopharyngeal area measurements for /mp/ in hamper are the best indicators of nasal emission. This speech sample should therefore be included in the assessment protocol. Inclusion of both 'pa' and 'pi' are useful in cases of variability of measurements, which may indicate variable velopharyngeal function.

7. 15 Summary of Pressure/Flow Results

This study developed a new approach to the assessment of nasal airflow errors in speech. The descriptive scale used four different categories of airflow errors and described them in terms of strength, consistency and frequency. No previous studies have evaluated perceptual measurements of nasal airflow in such depth, nor have they evaluated the relationship between nasal airflow errors and pressure/flow measurements.

The present findings indicated that pressure/flow measurements are valid assessments of nasal emission, but would need to be used in conjunction with perceptual assessment. Factors which influence this relationship will need to be considered in the clinical assessment of speakers with velopharyngeal dysfunction. Pressure/flow measurements were not valid assessments of nasal turbulence, nasal fricatives and velopharyngeal fricatives. The weak relationship between perceptual ratings of nasal turbulence and pressure/flow measurements underscores the need for further investigations into the acoustic and physiological correlates of nasal turbulence. If nasal turbulence is a type of velar trill, then a weak relationship between the measurements would be expected. As almost all incidences of nasal fricatives and velopharyngeal fricatives were phoneme specific, it is likely that they were not evident in the limited speech sample used in the pressure/flow measurements, resulting in a weak relationship between measurements.

Results indicated that nasal flow and velopharyngeal port area measurements are reliable and valid measurements of nasal emission. Although differential pressure measurements had good reliability and good validity for assessment of nasal emission, there was limited normative data available. This made it impossible to assess the efficiency of differential pressure measurements in distinguishing between normal and pathological speakers. A range of cut-off values was found to help distinguish between normal and pathological speakers using pressure/flow measurements. The data indicated that pressure/flow measurements during the /mp/ cluster in hamper were the best indicators of nasal emission.

7. 16 Limitations of the Study

Limitations of the Perceptual Profile have been outlined in Chapter 4 section 8.4. These include excessive detail of the profile and inadequate training of raters prior to the investigation. The number of speech parameters, coupled with an insufficient number of participants made it difficult to analyze results using the kappa reliability analysis. In order to evaluate the relationship between perceptual ratings and instrumental measurements, a value from 1 to 8 was allocated to represent the descriptive ratings of nasal airflow errors. For example, 1 represented weak, inconsistent and infrequent errors, while 4 represented strong inconsistent, and infrequent errors. This assignment of a value may have resulted in the loss of some information.

Only one nasal sentence was used in the speech stimulus for investigation of hyponasality. This was to ensure that the Total Test sentences had 11% nasal consonants, which is representative of the percentage of nasal consonants in normal speech (Fletcher et al., 1989). Although the present study found a good relationship between perceptual ratings of hyponasality and nasalance scores for the nasal sentence, the nasalance scores may not be reliable for assessment of individuals (Nichols, 1999). This weak reliability was noted in the test-retest analysis of the nasal sentence. It is recommended, from the present results and previous studies, that a separate speech stimulus be used to assess hyponasality. Nichols (1999) recommends that the stimulus should contain 10 nasal sentences.

The different speech stimuli used for the three assessments resulted in a limitation of the present study. Pressure/flow measurements were made during syllable and word production; however, perceptual measurements were made at the various levels of speech (i.e., syllable, sentence, automatic speech and conversation speech). Also, the perceptual assessment was not carried out simultaneously with instrumental assessment. This might have overcome some of the difficulties with inconsistent cases, where the level of nasality and nasal airflow errors may have varied between the two assessments.

A further limitation of the present study is that only a small and limited amount of data was analyzed for pressure/flow measurements. Because of the format of the PERCI

SARS, only the phoneme /p/ was used for measurements. This resulted in an inability of the system to measure nasal airflow errors in the contexts in which the most likely occur. The limited amount of data was necessary as the three assessments were carried out during the one assessment session in order to reduce the effects of variation in speech across time. It was necessary to keep the testing time short enough to ensure that the participant did not become fatigued.

The procedure for elicitation of the speech sample may have increased individual variability in sound production. Participants were asked to repeat the syllables 'pa pa pa' three times, and in many situations required prompting after each three repetitions. This type of elicitation procedure does not control for respiratory effort or variations in loudness and can result in increased individual variation in sound production. The standard procedure for recording data using the PERCI SARS is to elicit syllables and words in this way. This ensures that at least 6 peak oral pressure points can be identified and measured during a 10 second recording. It was hoped that by using mean pressure/flow measurements, the influence of individual variation in sound production could be minimized; however, the test-retest results of the pressure/flow measurements indicate that this was not the case. The use of a short carrier phrase to elicit the word 'hamper' may have minimized the effects of respiratory effort and increased loudness.

7. 17 Summary

Results of perceptual assessments indicated that almost half the population studied presented with mild speech problems. There was however, a complete range of speech errors was included in the study population. The data indicated that the Perceptual Profile, Nasometry (using the Total Test sentences and HP consonant sentences), and Nasal Flow and Velopharyngeal Port Area measurements were reliable assessments of nasality and nasal airflow errors in speech. Validity of the profile has been shown, as well as validity of the combined approach to assessment using perceptual, nasometric and nasal flow/velopharyngeal port area measurements. Results of the relationship between perceptual and nasometric measurements in the present study compared well with previous studies. The relationship between pressure/ flow measurements and perceptual measures varied considerably. However, pressure/flow measurements as an indicator of nasal emission produced promising results.

The issue of cut-off values for instrumental measurements was discussed. Results indicated that a range of values, from statistically determined values to clinically determined values, was useful in identifying patients with pathological speech. The use of different speech stimuli in instrumental assessments was evaluated.

Factors that influenced the relationship between instrumental and perceptual measurements were identified. These include: consistency of nasality and/or nasal airflow errors, associated nasal airflow errors, voice problems and articulatory errors, errors in timing of velopharyngeal closure, inaudible nasal emission, and increase in respiratory and articulatory effort These factors should be considered in clinical assessments, as they will significantly influence the type of management required and the outcome of this management.

Conclusions and recommendations from the above results are presented in the following chapter.

CHAPTER 8

CONCLUSIONS

Introduction

The primary aim of the present study was to develop a reliable and valid perceptual assessment scale, which could be used by clinicians for evaluating nasality and nasal airflow problems in speech. This involved the development of a descriptive Perceptual Profile (Chapter 4) and subsequent evaluation of its reliability (Section 8.1). In order to validate the Perceptual Profile, instrumental assessment techniques had to be identified and evaluated. The second aim of this study was to develop a combined perceptual and instrumental protocol for specialized assessment of nasality and nasal airflow problems, which are usually undertaken in a centre of expertise for Cleft Lip and Palate. The present study investigated the ability of the instruments to distinguish between normal and pathological speakers (Section 8.2). Perceptual, Nasometric and Pressure/Flow measurements were validated by assessing the relationship between the measurements (Section 8.3). A discussion of the clinical implications of the study (Section 8. 4) and recommendations for further research (Section 8. 5) are presented in this chapter.

8. 1 Perceptual Profile

A descriptive perceptual scale for the assessment of nasality and nasal airflow was developed for the present study. Problems of perceptual assessments, such as inadequate definitions of parameters, reproducibility of test results, and diagnostic classification were addressed in the development of a Perceptual Profile. The profile was based on definitions of nasality and nasal airflow allowing for separate ratings of nasality and nasal airflow categories. This profile represents a new approach to assessing nasality and nasal airflow problems, providing a descriptive rather than a quantitative scale. Initial reliability studies indicated that:

1. The Perceptual Profile had good intra-rater reliability for the author, but these results could not be generalised to other raters. There was good reproducibility of

test results, indicating adequate reliability for comparison with instrumental assessments. Analysis of listening conditions indicated that there were differences in reliability of ratings of nasal turbulence, nasal fricatives and velopharyngeal fricatives between live and audio assessments. Previous studies of listening conditions (Moller & Starr, 1984) had not evaluated assessment of different types of airflow errors.

2. Inter-rater reliability of the Perceptual Profile was inconclusive, with a good agreement, but poor Kappa scores. Despite this, the inter-rater agreement for ratings of nasality, using the Perceptual Profile, compared well with inter-rater agreement in previous studies (Sweeney, 1984; Sell & Grunwell, 1990; Sell et al., 1994; Lohmander-Agerskov, 1996; Watterson et al., 1998; CSAG, 1998). The present results compared well with previous studies of the reliability of nasal emission (Sell et al., 1994; Lohmander-Agerskov, 1996; CSAG, 1998). No other study assessed reliability of other nasal airflow errors. The results of the present study raise serious questions regarding previous reliability studies and indicate the need for rigorous evaluation of reliability of perceptual assessments of nasality and nasal airflow errors in speech.

It was evident from the literature that there was no one ideal measurement of nasality and nasal airflow. Therefore, the critical question was to identify which combination of different assessments would provide the most insights about any particular experimental work or clinical question (Folkins & Moon, 1990). The present perceptual results underlined the need to supplement perceptual assessment of nasality and nasal airflow errors with instrumental measurements.

8. 2 Nasometry

Previous studies have indicated that the Nasometer was a useful assessment tool for investigation of nasality, when selected speech stimuli were used. In the present study, careful consideration was given to the speech stimulus used for nasometry. It was intended that the same speech stimulus be used for perceptual assessments. As the speech stimulus was to be used for both assessments, it was important that the stimulus contained all speech consonants and a normal percentage of nasal consonants. However, it was also necessary to have a speech stimulus that contained high pressure consonants, which was devoid of nasal consonants, and a stimulus that contained low pressure consonants, also devoid of nasal consonants. Previous studies have found that the nasalance scores on speech stimuli containing high pressure consonants and devoid of nasal consonants were good indicators of hypernasality (Dalston et al., 1991b; Vallino-Napoli & Montgomery, 1997). Karnell (1995) had stated that separate high pressure consonant and low pressure consonant speech stimuli should be used to assess the effect of nasal airflow errors on nasalance scores. The speech stimulus used in the present investigation allowed for separate measurements of nasalance scores on sentences containing all consonant types: High Pressure consonant sentences only, Low Pressure consonant sentences and a Nasal sentence. Conclusions indicated that:

1. The Nasometer did distinguish between normal speakers and speakers with velopharyngeal dysfunction. The Total Test sentences, the sub categories of High Pressure consonant sentences and Low Pressure consonant sentences distinguished between normal speakers and speakers with hypernasality. The Nasal sentence distinguished between normal speakers and speakers with hyponasality.

2. Contrary to the findings of Karnell (1995), participants with nasal turbulence, or nasal emission, did not have greater nasalance scores on High Pressure consonant sentences compared to Low Pressure consonant sentences. The data indicated no benefit in the use of separate nasalance scores for High Pressure and Low Pressure consonant sentences.

3. Test-retest analysis indicated that the speech stimulus containing all consonant types was reliable and the sub-categories of High Pressure consonant and Low Pressure consonant sentences were also reliable. However, the single nasal sentence had weak reliability.

Comparison of perceptual ratings of nasality and nasometry indicated the following:

1. Sentences which contained high pressure consonants and were devoid of nasal consonants had a strong relationship with perceptual judgments of nasality. This supports the results of previous studies (Dalston et al., 1991b; Dalston & Seaver 1992). The results indicate that the High Pressure consonant sentences are valid speech stimuli for the nasometric assessment of hypernasality.

2. However, contrary to previous studies, the present study indicated a good relationship between perceptual ratings of hypernasality and the Total Test sentences. These sentences were used as they contained all consonant types, as well as 11% nasal consonants, which is similar to the proportion of nasal consonants in conversational speech (Fletcher et al., 1989). The improved relationship may be due to the descriptive scale used in the present study. The present results indicated that the Total Test sentences are valid speech stimuli for the nasometric assessment of hypernasality. However, due to the fact that silent intervals can influence nasometry scores, it is recommended that the silent gaps between each sentence category be deleted when calculating the nasalance scores for the Total Test sentences.

3. Despite weak test-retest reliability for the nasal sentence, this sentence was a valid measurement of hyponasality. However, due to poor reliability of the nasal sentence further investigation into nasometry and hyponasality is recommended for clinical use.

4. When there is a discrepancy between perceptual and nasometric results, factors which can influence the relationship between measurements should be considered and investigated.

5. Results indicated that a combination of statistically determined and clinically determined cut-off values produce best overall efficiency rating for nasometry. It was concluded that a range of cut-off values was useful in distinguishing between normal speakers and speakers with velopharyngeal dysfunction (Table 8.1).

Table 8.1 Optimal cut-off nasalance values for each speech category for distinction between normal and nasal speakers.

	Total Test	**High Pressure**	**Low Pressure**	**Nasal**
Optimal cut-off values	**30% - 35%**	**24% - 29%**	**23% - 28%**	**37% - 42%**

This new descriptive Perceptual Profile, along with nasometry measurements for the Total Test sentences and High Pressure consonant sentences, provides a valid protocol for the assessment of hypernasality.

8. 3 Pressure/Flow Measurements

Conclusions of the series of investigations of the pressure/flow measurements in the present study indicated the following:

1. Nasal flow measurements and velopharyngeal port area measurements were able to distinguish between normal speakers and speakers with velopharyngeal dysfunction.

2. Oral pressure measurements on /p/ did not differ sufficiently to allow an assessor to distinguish reliably between normal speakers and speakers with velopharyngeal dysfunction.

Comparison of perceptual ratings of nasal airflow and pressure/flow measurements indicated that:

1. Nasal flow and velopharyngeal area measurements have good relationships with perceptual ratings of nasal emission. The best results were obtained from pressure/flow measurements during the production of /p /in hamper. Although results were good, correlations and overall efficiency measures were not high enough to suggest that pressure/flow measurements, on their own, are valid and reliable measures of nasal emission. However, when used in conjunction with perceptual measurements, nasal flow and velopharyngeal area measurements have been found to be useful for assessment of nasal emission.

2. Oral pressure measurements were not good indicators of nasal airflow errors and, therefore, were considered inadequate for assessment of these types of errors.

3. Differential pressures provided a good indication of nasal emission only for the speech sample 'hamper'. Insufficient normative data was available for differential pressure measurements to date. Further investigation into differential pressure measurements is recommended.

4. None of the pressure/flow measurements related well with perceptual ratings of nasal turbulence, nasal fricatives and velopharyngeal fricative. Hence, these instrumental measurements, used in this study, were not valid measurements of nasal airflow errors other than nasal emission.

5. When discrepancies occur between perceptual and pressure/flow measurements, factors which influence the relationship between measurements need to be considered and investigated.

6. A combination of statistically and clinically determined cut-off values were found to be useful for clinical evaluation of nasal emission. As a result, a range of cut-off scores is recommended for assessment (Table 8.2).

Table 8.2 Optimal cut-off values for pressure/flow measures in the identification of nasal emission.

	Nasal Flow			Velopharyngeal Port Area		
	pa	**pi**	**hamper**	**pa**	**pi**	**hamper**
cut-off values	4 - 9 ml/s	4.4 ml/s	14.5 - 30 ml/s	.004 - .01 cm^2	.002 cm^2	.007 cm^2

It can be concluded that the instrumental measures of nasal flow and velopharyngeal port area and perceptual ratings of nasal emission provide a valid assessment protocol for the evaluation of nasal emission in speakers with velopharyngeal dysfunction. However, these results reveal the need to adapt the Perceptual Profile for the assessment of nasal turbulence, nasal fricatives and velopharyngeal fricatives. Further investigation into the instrumental assessment of these airflow errors is recommended.

8. 4 Recommended Changes to the Perceptual Profile

It is evident from the reliability studies that the present Perceptual Profile is too detailed, and some revisions are required. However, authorities have suggested that although reducing the detail of a scale improves overall reliability, this may result in loss of information (Henningsson & Hutters, 1997). This investigation has identified the sections of the Perceptual Profile that require revision. The nasality section has been shown to have an acceptable level of reliability and validity in the evaluation of nasality. Hence, no revisions of the nasality assessment are required. However, the results of the reliability and validity studies of the nasal airflow sections suggested the need for revision. In the section on nasal emission and nasal turbulence, it is recommended that phoneme specificity be omitted. This parameter had the weakest agreement on the intra- and inter-rater reliability studies, possibly due to the use of the term phoneme specific. In the literature and in clinical evaluation, phoneme specific usually refers to a nasal airflow error that substitutes a phoneme (Trost, 1981; Albery, 1989; Sell et al., 1994;1999). The reliability data indicates that raters did not distinguish between the categories of nasal fricatives and velopharyngeal fricatives. Results indicate the need to combine these airflow errors into one category. It is recommended that these errors will be recorded as present or absent, and if they are phoneme specific, this will be recorded. The revised Perceptual Profile is presented below (Figure 8.1).

Name: **Age:** **Date:**

NASALITY / NASAL AIRFLOW ASSESSMENT

Nasality

Hypernasality: Present Absent
a) mild, evident but acceptable.
b) mild/moderate, unacceptable distortion , evident on close vowels
c) moderate, evident on close and open vowels.
d) moderate / severe, evident on all vowels and some consonants.
e) severe, evident on all vowels and most voiced consonants.
Consistent Inconsistent

Hyponasality: Present Absent
a) evident, but acceptable
b) moderate - all vowels reduced nasality and some nasal consonants.
c) severe - total denasal production of nasal consonants.
Consistent Inconsistent

Nasal Airflow

Nasal Emission:	Strong	Weak
	Frequent	Infrequent
	Consistent	Inconsistent
Nasal Turbulence:	Strong	Weak
	Frequent	Infrequent
	Consistent	Inconsistent
Nasal / Velopharyngeal Fricative	Present	Absent
	Phoneme specific Y/N	

____

Test Words
Articulation ______________________________________

Nasality ______________________________________

Sentence Repetition
Articulation ______________________________________

Nasality ______________________________________

Automatic Speech

Jack / Jill ______________________________________

Counting 1 - 20 __________________ 60 -70 ______________

Conversational Speech

Articulation ______________________________________

Nasality ______________________________________

Figure 8. 1 Revised Perceptual Profile.

These data suggest that there is a need to undergo considerably in depth training in the use of the Perceptual Profile. It is recommended that future training will include the following:

- definitions and audio taped examples of terms used to describe nasality and nasal airflow errors in speech. Details of the terminology presented in Chapter 4 will be included;
- explanation and audio taped examples of the descriptive scale used in the Perceptual Profile, ensuring that the listener understands the descriptive concepts such as weak/strong, consistent/inconsistent and frequency;
- group rating of audio taped speech samples, where the perceptual rating of each speech parameter is discussed and agreed;
- on completion of training, listeners will rate audio recorded speech samples and reliability of ratings will be evaluated.

The Nasometer should be used to supplement perceptual measures of nasality. Nasalance scores for the Total Test sentences and High Pressure consonant sentences can be used as indicators of hypernasality. When perceptual ratings of nasality and nasalance scores agree, the assessor can have confidence in the qualitative and quantitative measurements. When there is a discrepancy between the two measurements, possible causes of the difference in measurements needs to be examined. Further investigations may be required to examine the extent of factors such as consistency of nasality, associated nasal airflow errors and borderline velopharyngeal functioning, and to examine how these factors may determine future management.

Pressure/flow measurements, specifically nasal flow and velopharyngeal port area measurements, provide a useful and quantitative supplement to the perceptual measures of nasal emission. Possible causes for discrepancies in assessment results need to be examined when there are differences in perceptual and pressure/flow measurements. Consistency of nasal airflow errors and/or timing errors of velopharyngeal closure will require further investigation. If there are compensatory changes within the vocal tract, these changes will need to be evaluated and interpreted accordingly. Any change in vocal behaviour will have significant clinical implications and will require management.

Speakers with inaudible nasal emission will require monitoring with facial growth and adenoidal hypertrophy. Such speakers may be at risk for development of velopharyngeal inadequacy with growth and maxillary advancement.

The use of the three assessment techniques in the evaluation of the speech of individuals with cleft palate, velopharyngeal dysfunction or nasal obstruction is presented in Figure 8.2. This flowchart is adapted from a scientific approach to the assessment of voice disorders, developed by Behrman and Orlikoff (1997). A hypothesis is formulated about the nature, severity and possible causes of the problem, based on information from a number of sources, including case history, medical data and perceptual assessments. The hypotheses are then tested using instrumental measures. If the hypotheses are supported by the instrumental measures, further investigations using nasendoscopy and videofluoroscopy can be carried out, or therapy may be recommended. Direct observation may then indicate the need for surgery, prosthetics, therapy or a combination. Following treatment, reassessment will be required. If the hypotheses are not supported by the instrumental results, factors such as consistency of nasality, associated nasal airflow errors and borderline velopharyngeal functioning, which influence the relationship between perceptual and instrumental measurements need to be examined and the relevance of such factors needs to be ascertained.

In conclusion the descriptive scale of nasality and nasal airflow, used in the present study, provides an assessment tool that has good intra-rater reliability for the author, acceptable inter-rater agreement and good validity. Variable inter-rater reliability emphasizes the need to supplement the perceptual assessment with instrumental assessment. These data indicate that a combined perceptual and instrumental assessment protocol provides a valid and reliable assessment of nasality and nasal emission. When both perceptual and instrumental measurements agree, the examiner can have confidence that his/her findings provide a valid, reliable measurement of speech. When the perceptual and instrumental measurements disagree, the examiner needs to investigate causes of discrepancies in results.

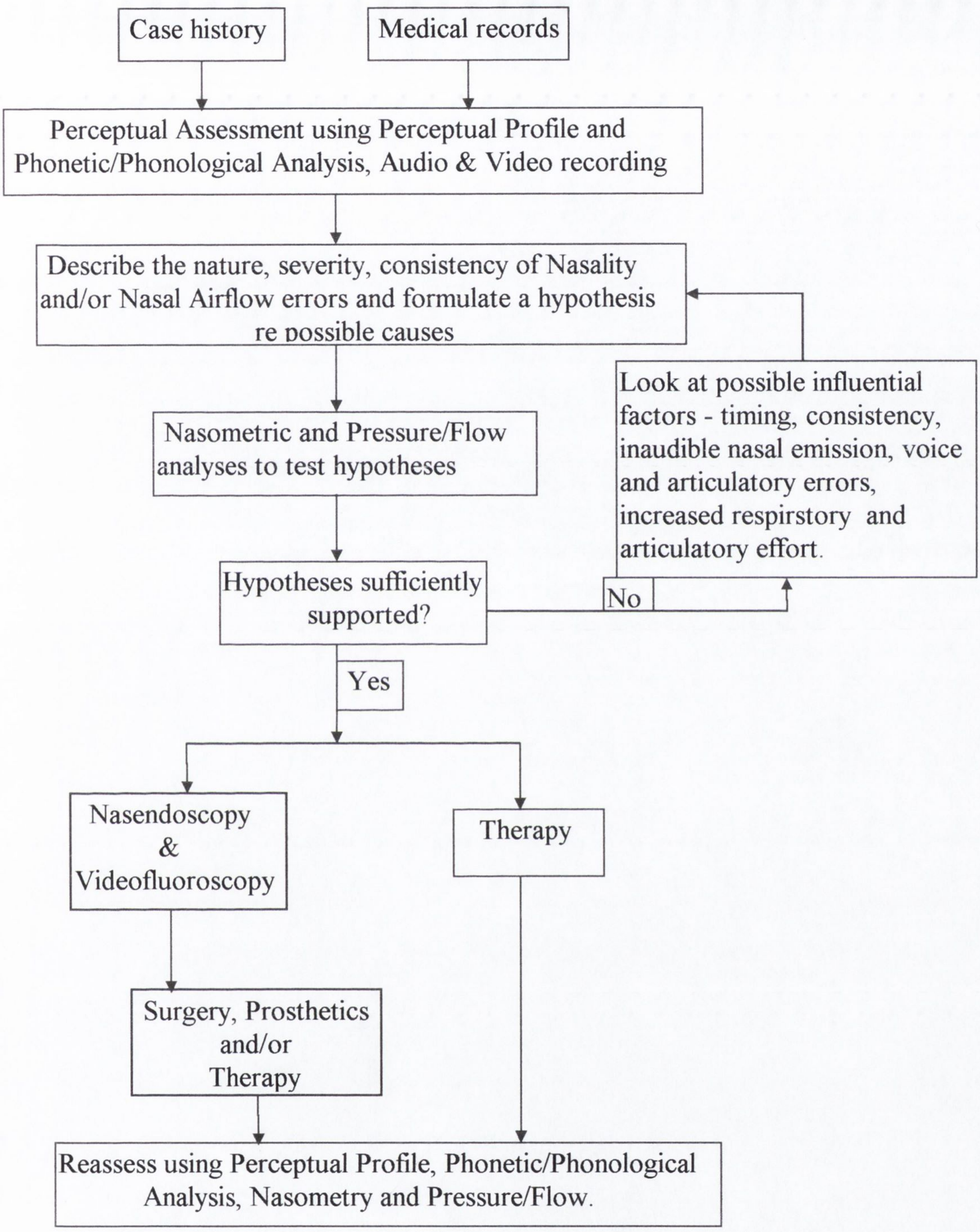


Figure 8.2 Application of combined perceptual and instrumental assessment of nasality and nasal airflow errors in speech. Adapted from Behrman and Orlikoff (1997).

8. 5 Recommendations for Further Research

- Further investigation into the reliability of the revised Perceptual Profile is recommended. Intra- and inter-rater reliability needs to be assessed using different raters. This will require in depth training on the use of the Perceptual Profile. Auditory training, to help identification of the nasal airflow categories, would be required. Definitions of terms, including definition of consistency and frequency, will need to be outlined. In order to use a kappa analysis of reliability data, a larger study population will be required.

- The use of the Nasometer in the assessment of hyponasality requires further investigation. A larger speech sample, with a high percentage of nasal consonants, should be used. Hence, a separate speech stimulus would be developed for nasometric assessment of hyponasality, and normative data will have to be obtained. Further investigation into the relationship between perceptual assessments of hyponasality and nasometric assessment using a speech sample with a high percent of nasal consonants, is recommended.

- Further investigation into the use of pressure/flow measurements for the assessment of nasal turbulence and nasal/velopharyngeal fricatives is recommended. A study of the validity of the pressure/flow systems using a Rothenberg mask, instead of the tube system, is recommended. The Rothenberg mask is a divided mask with a nasal section and a separate oral section. Each section contains a pressure and flow transducer, resulting in measurements of nasal flow and nasal pressure, and oral flow and oral pressure. The mask would allow for assessment of larger speech samples using a range of consonants. Investigations into the reliability of the system using the mask is recommended.

- Further studies on the relationship between perceptual ratings of nasal emission and pressure/flow measurements, while controlling for loudness, articulatory and respiratory effort are recommended. This could be carried out by eliciting the word 'hamper' in a carrier phrase for the pressure/flow measurements.

- Errors of timing of velopharyngeal closure were evident in some participants who had poor correlations between perceptual measures of nasal airflow and pressure/flow measurements. Timing errors can be identified by measuring the timing relationship between nasal flow measurements and oral pressure measurements. Further investigation into the incidence and type of timing errors is indicated.

- Investigations into the articulatory production of nasal turbulence and velopharyngeal fricatives is recommended in order to ascertain if the sound is a result of narrowing of the velopharyngeal port or rapid contact between the velum and pharyngeal walls.

- The acoustic correlates of the nasal airflow errors requires further investigation. Preliminary results indicate that there is low frequency energy associated with nasal turbulence/velopharyngeal fricatives, but this requires more detailed spectral analysis.

- The presence of silent intervals in nasometric measurements was found to produce high nasometric scores in analysis of some speakers. The effect of silent intervals on overall nasalance scores needs to be investigated, and possible explanations postulated.

- Variations in compensatory vocal tract needs to be investigated to ascertain the types of changes speakers adopt when velopharyngeal inadequacy is present and the influence of these changes on perceptual and instrumental assessments.

Appendix 1

Glossary

Acoustic assessment involves the measurement of acoustic energy during sound production

Aerodynamic assessment is the measurement of airflow volumes and pressure flows during speech

Agreement implies that two listeners assign identical meanings to each scale point

Cleft of the palate, is an opening in the palate resulting in the continuous passage between the mouth and nose, where the palate serves as the roof of the mouth and the floor of the nose

Compensatory articulation errors are a distinct category of errors commonly found in patients with velopharyngeal dysfunction. Sounds are produced as compensatory articulations producing abnormal articulation at a place where success is more likely such as glottal or pharyngeal articulation

Cut-off values indicates a measurement above which speech can be considered abnormal

Differential pressure is a measurement of the difference between oral pressure and nasal pressure

Dysphonia refers to a disorder of voice at the level of the larynx

Hypernasality is an increase of perceived nasal resonance of the voice

Hyponasality is the perceived reduction in normal nasal resonance resulting from a blockage of the nasal airway (D'Antonio & Scherer, 1995)

Instrumental assessment refers to computer based assessments of the acoustic and aerodynamic correlates of speech

Level of speech refers to the levels of complexity of speech from word to sentence to automatic and conversational speech

Nasal airflow errors is a generic term describing the inappropriate escape of air through the nasal cavity during production of voiced and voiceless oral pressure consonants, and is usually associated with incomplete velopharyngeal closure

Nasal emission is audible escape of air through the nasal cavity accompanying production of oral pressure consonants

Nasal obstruction is a blockage within the nasal cavity, which may prohibit the normal production of nasal consonants and possibly influence nasal airflow

Nasal pressure is a measurement of air pressure within the nose during sound production.

A *nasal fricative* has frictional sound produced by air passing through the nasal cavity when there is incomplete velopharyngeal closure and no audible oral release of air.

Nasal turbulence is defined as a 'snorting' or turbulent noise resulting from the approximation but inadequate closure of the superior border of the velum and the posterior pharyngeal wall.

Nasalance is a numeric ratio of nasal acoustic energy to the sum of nasal plus oral acoustic energy, multiplied by 100.

Nasality describes the perceptual suprasegmental feature associated with the acoustic response of air within the coupled oral and nasal cavity.

Nasometer is a computer-based instrument which measures nasal and oral acoustic energy during speech.

Oral pressure is a measurement of pressure within the oral cavity during sound production

Perceptual assessment refers to listener based judgements of speech

PERCI SARS (Speech Aero-Dynamic Research System) is a computer software/hardware interface for in-depth assessment of speech aeromechanics and nasal airway patency

Pharyngeal flap is an operation, which is carried out in order to reduce the size of the velopharyngeal port. A flap from the posterior pharyngeal wall and the velum are attached to form a bridge in the central of the port

Reliability refers to a mathematical relationship between an observed test score and the test takers 'true score' on the test

Sentence category is the categorisation of sentences according to consonant type

Test sensitivity is the percentage of participants who are identified as having abnormal speech on one test and who are also identified as having abnormal speech on another test.

Test specificity refers to the percentage of patients who are identified as having normal speech on one test and who are also identified as having normal speech on another test.

Validity refers to whether measurements actually measure what they are designed to measure.

Velopharyngeal dysfunction is a generic term, which denotes any type of abnormal velopharyngeal function resulting from structural deficits, neurological disorders, faulty learning or a combination of aetiologies

A *velopharyngeal fricative* is a 'snorting' or turbulent noise, resulting from the approximation but inadequate closure of the superior border of the velum and the posterior pharyngeal wall, and functions as a substitution for another sound

Velopharyngeal port area is an estimated measurement of the velopharyngeal port calculated using an equation of pressure difference between the nose and mouth and airflow through the nose

Appendix 2

Pittsburgh Sentences

1. Mama made lemon jam.
2. Put the baby in the buggy.
3. Kindly give Kate the cake.
4. Go get the wagon for the girl.
5. Sissy sees the sun in the sky.
6. The ship goes in shallow water.
7. Jim and Charlie chew gum.
8. Please tie the stamps with string.

(Phillips, 1986)

GOS. SP. ASS Sentences.

Mum came home early.
The puppy is playing with the rope.
Bob is a baby boy.
The phone fell off the shelf.
Dave is driving a van.
Neil saw a robin in a nest.
A ball is like a balloon.
Tim is putting a hat on.
Daddy mended a door.
I saw Sam sitting on a bus.
The Zebra was at the zoo.
Sean is washing a dirty dish.
Charlie's watching a football match.
John's got a magic badge.
The bell's ringing.
Karen is making a cake.
Gary's got a bag of Lego.
Hannah hurt her hand.
This hand is cleaner than the other.
The hamster scrambled up Stuart's sleeve.

(Sell et al., 1999)

Appendix 3

Standard passages provided by the manufacturers of the Nasometer

Rainbow Passage

When sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colours. These take the shape of a long round arch, with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

(Fairbanks, 1960)

Zoo Passage

Look at this book with us. It's a story about a zoo. That is where bears go. Today it's very cold out of doors, but we see a cloud overhead that's a pretty white fluffy shape. We hear that straw covers the floor of cages to keep the chill away; yet a deer walks through the trees with her head held high. They feed seeds to the birds so they're able to fly.

(Fletcher, 1972)

Nasal Passage

Mama made some lemon jam.

Ten men came in when Jane rang.

Dan's gang changed my mind.

Ben can't plan on a lengthy reign.

Amanda came from Bounding, Maine.

(Fletcher,1972)

Appendix 4.

<u>Turtle passage</u>

What could it be?
It has a head,
Four feet with a tail,
It walks real slow,
Cause it carries a house,
Could it be turtle?

<u>Mouse passage</u>

There was a young mouse
That lived in my house
Who wanted to go to school.
He asked one day,
Please show me the way,
I said yes, follow me,
Mr Mouse was going to school.

(Watterson et al., 1996)

Appendix 5

Bobby and Billy Play Ball

Bobby and Billy go to play ball.
They get a bat, a ball and a glove.
They go to the ballpark.
Bobby took a turn at bat.
Bobby tried to throw the ball.
Billy hit the ball up high.
Bobby and Billy like to play ball.

A school day for Suzy

Suzy eats cereal or toast for breakfast.
After that, she rides the bus to school.
Susie likes to sit with Sally.
At school, the teacher gives Suzy's class a test.
Suzy likes her school.
She also likes teacher.

(MacKay & Kummer, 1994)

Appendix 6

Nasalance scores for normal Irish English-speakers - Pilot Studies

Pilot Study 1

Participants

Twenty one children with the Irish-English accents aged between 5 and 10 years were assessed. The children attended an orthopaedic outpatient department of The Children's Hospital. All were judged to have normal resonance and airflow either by an experienced Speech and Language Therapist or by an experienced ENT surgeon. Children were excluded if they had a history of speech or palatal problems, an upper respiratory tract infection or hearing loss on the day of assessment.

Instrumentation

The Nasometer (Model 6200.2) was used in the initial pilot study. The headset was placed on the child's head in the appropriate position. (i.e. with the sound separator below the nose and above the upper lip without interfering with lip movement)

Speech sample

1. the 10 test words included in the perceptual assessment profile
2. 15 test sentences
3. A nursery rhyme (*Jack and Jill*)
4. Four conversation type sentences e.g. *My name is....., I live at.......*

   [Appendix 10]

Method

Each child repeated the speech sample. The child's responses were saved in the analysis section of the Nasometer. All speech stimuli were recorded on a 20 second time display. The words and nursery rhyme were analysed on the 20 second time display, while the sentences and conversational speech were analysed on a 40 second time display.

An overall nasalance score was calculated for the complete speech sample. Cursors were placed at the beginning and the end of each of the following sections of the speech sample: words; sentences; automatic speech and conversational speech. Hence, separate nasalance scores were calculated for words, sentences, automatic speech and conversational speech.

Results of Pilot Study

Nasalance scores were calculated initially for the complete speech sample, including words, sentences, automatic speech and conversational speech (Figure 1). Results indicated a mean nasalance score of 30%, a standard deviation score of 10%. Scores for the entire speech sample ranged from 11% to 56%.

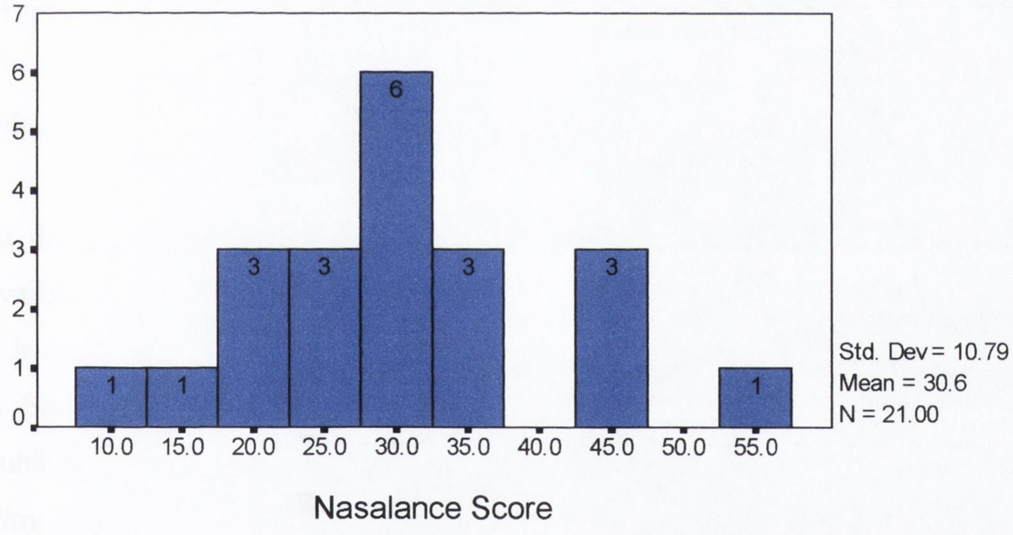


Figure 1. Mean nasalance score, standard deviation and distribution of nasalance scores for 21 normal speakers during the production of words, sentences, automatic speech and conversational speech.

Analysis of separate levels of speech, i.e. word, sentence, automatic and conversational speech, proved to be extremely difficult. It was not possible to identify the beginning and the end of the speech level on the nasalance graph. Only the sentence level could be reliably identified on the nasometry graph. Insufficient space had been left between each type of speech sample. (i.e. between words and sentences, between sentences and automatic speech, and between automatic speech and conversational speech)

Nasalance scores for the sentences were reliably identified in 18 of the 21 cases. Results indicated that scores ranged from 19% to 35% with a mean of 28%.

Discussion

Results indicated a mean nasalance score for normal speakers of 30%, ranging from 11% to 56%. Previous American studies indicated normal nasalance means of approximately 36% (Fletcher et al., 1989; Seaver et al., 1991). However, comparison with the previous studies was difficult due to the use of different and longer speech stimuli in the present study. Analysis of the nasalance data indicated that the speech sample was too large and cumbersome to allow for reliable identification of different levels of speech and hence, no reliable nasalance values were obtained for these levels of speech. A further pilot study was recommended using a shorter speech sample.

Pilot study 2

At the time of the second pilot study, an updated version of the Nasometer became available. The updated nasometry programme (model 6200-3) was used in this study. It included an improved menu format, improved graphics and a simpler analysis format. A Simplified Nasometric Assessment Procedure (MacKay & Kummer, 1994) was published with the upgraded Nasometer manual. A short passage, the *Bobby and Billy Play Ball* passage, provided normative nasalance scores for American children. This passage was included in the second pilot study to compare Irish-English nasalance score with American English nasalance scores. The original test sentences were adapted to include sentences which contained low pressure consonant sentences only.

Participants

Further data was collected on 26 children between the ages of 3.6 to 12 years. All children had Irish-English accents and were judged by experienced Speech & Language Therapist to have normal speech .

Speech Sample

1. The sentence test sample was changed to include two sentences containing only low pressure consonants. One sentence containing /s/ and /n/ phonemes was deleted, as these phonemes were already included in the sample. The 16 test sentences were presented in the following order: high pressure consonant sentences; low pressure consonant sentences; mixed consonant sentences and a nasal sentence.
(Appendix 7)

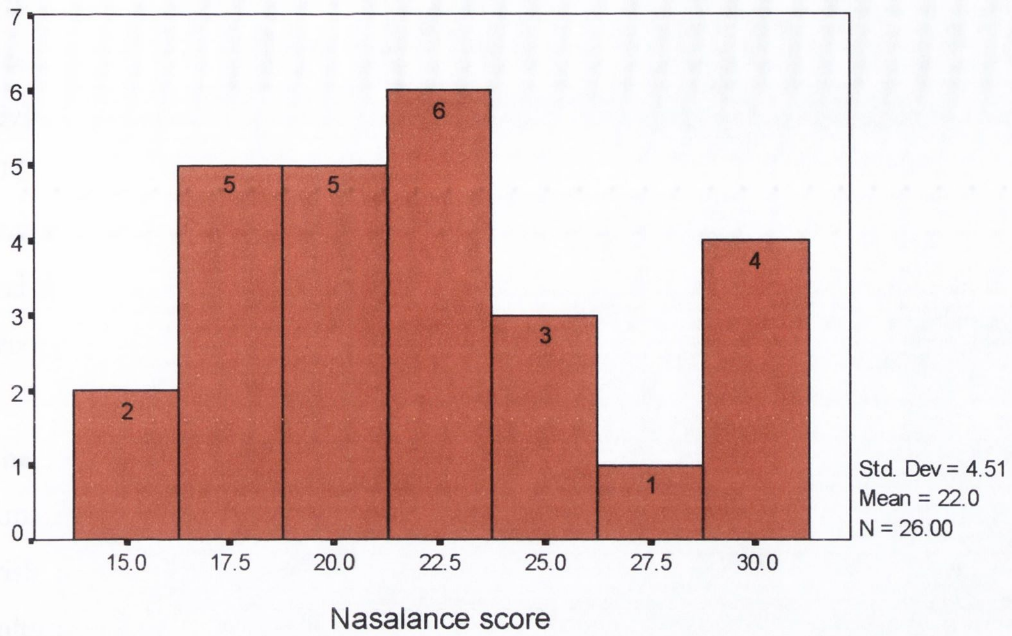
2. The *Bobby and Billy Play Ball* passage, which is a subtest of the Simplified Nasometric Assessment Procedure (MacKay & Kummer, 1994). (Appendix 5)

Method

The headset was placed in the appropriate position on the child's head with the sound separator above the upper lip and below the nose. The child repeated 16 test sentences. The sentences were recorded at a time axis of 4 seconds with a one second gap between each sentence. A 5 second gap was left between the sentences and the *Bobby & Billy* passage. Analysis was made at 20 second display to allow easy identification of the beginning and end of the two speech samples. In order to record the *Bobby & Billy* passage, the child was asked to repeat each line of the passage after the examiner.

Results.

A mean nasalance score of 22% was found for 26 normal speakers during the production of the 16 test sentences. A standard deviation of 4.5% was found, with a range of 15% to 30% (Figure 2).



Analysis of the nasalance scores for the *Bobby and Billy* passage indicated a mean nasalance score of 14.6%, a standard deviation score of 4.8% and a range of 6% to 24% (Figure 3).

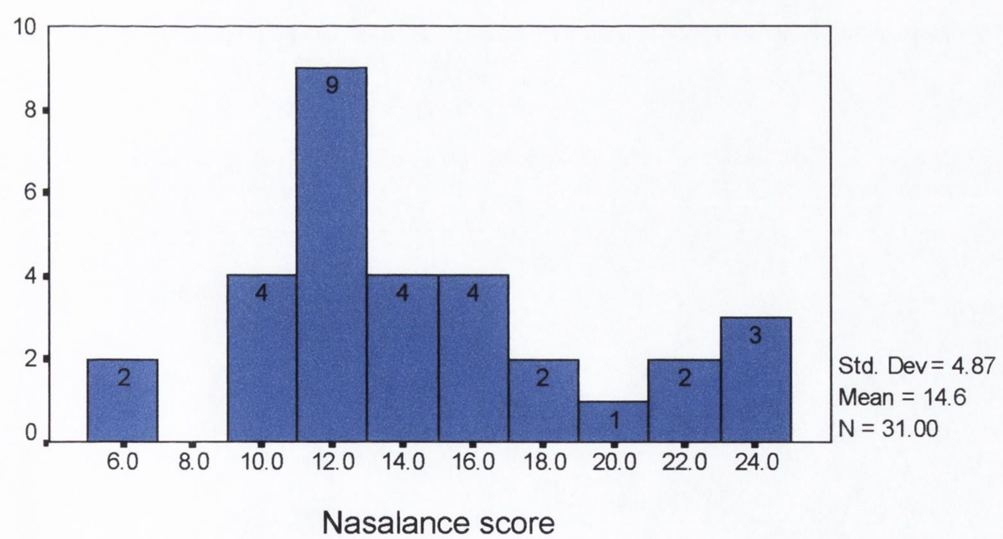


Figure 3. Mean and distribution of nasalance scores for 26 normal speakers during production of the '*Bobby and Billy'* passage.

Discussion

Results of the second pilot study indicated a lower mean nasalance score for the revised test sentences compared to the mean nasalance score for the original sentences. This was probably due to the change in the speech stimulus. Analysis of the separate sentence categories was not possible as it was not possible to identify the beginning and the end of each sentence category. It was recommended that in future studies a 5 second gap should be left between sentence categories.

The mean nasalance scores on the *Bobby and Billy* compared favourably to the American data. The American mean is 15.4% (MacKay & Kummer, 1994). However, Irish children found it difficult to repeat the *Bobby and Billy* passage due to the cultural difference in meaning. Irish children would not bring a bat or a glove to play ball and many of the participants omitted, repeated or hesitated during repetition of the passage. It was concluded that the *Bobby & Billy* passage was unsuitable for further nasalance studies on Irish-English speakers.

Appendix 7

<u>**Test sentences**</u>

1.	**Paul likes apple pie**	
2.	**Gary's got a bag of lego**	
3.	**Vickey's got a very heavy bag**	**High pressure**
4.	**The zebra lives at the zoo**	**consonants**
5.	**The shoe shop was shut**	
6.	**We were away all year**	**Low pressure**
7.	**Will you wear a lily**	**consonants**
8.	**Ben is a baby boy**	
9.	**Tim had a tart for tea**	
10.	**Daddy mended the door**	**Mixed**
11.	**Kevin's looking at the book**	**consonants**
12.	**The phone fell off the shelf**	
13.	**I saw Sam sitting on a bus**	
14.	**John jumped off the bridge**	
15.	**The children were watching a football match**	
16.	**Mum came home early**	**Nasal**
		consonants

Appendix 8

Nasometry Test-Retest Reliability

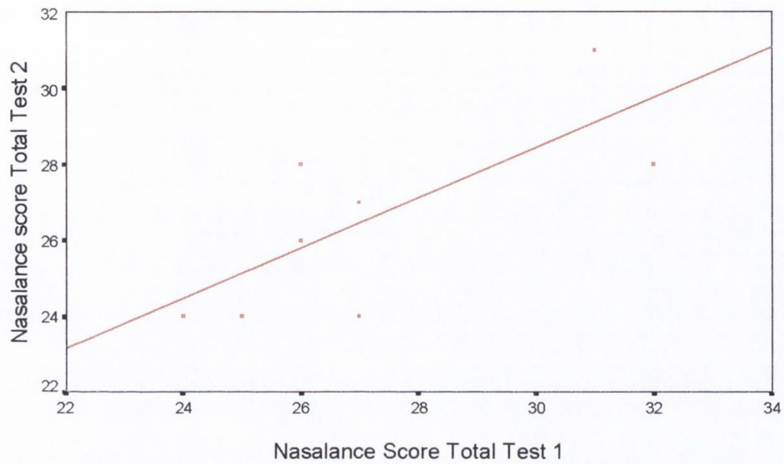


Figure 8a . Test-retest relationship for nasalance scores during the repetition of the Total Test Sentences by 10 normal speakers.

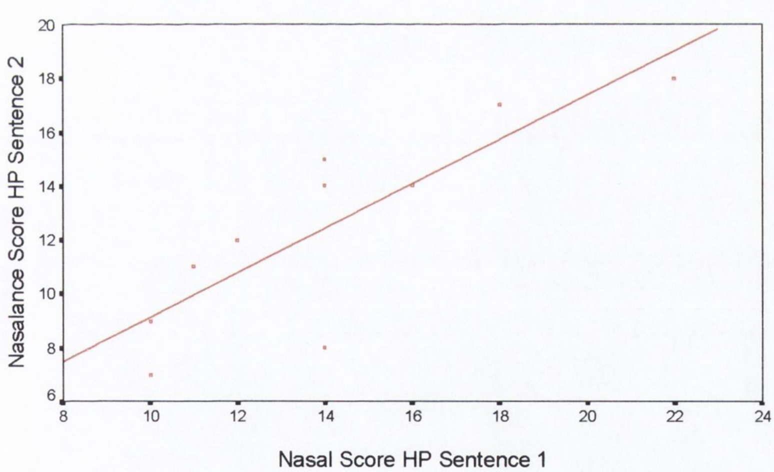


Figure 8b . Test-retest relationship for nasalance scores during the repetition of High Pressure Consonant Sentences by 10 normal speakers.

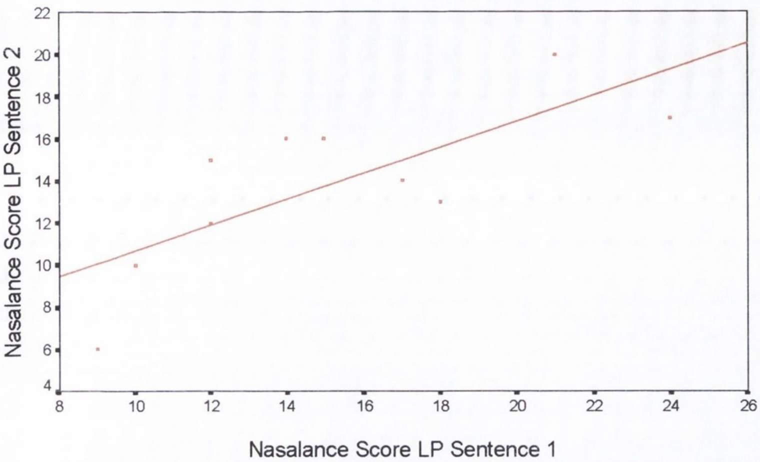


Figure 8c . Test-retest relationship for nasalance scores during the repetition of Low Pressure Consonant Sentences by 10 normal speakers.

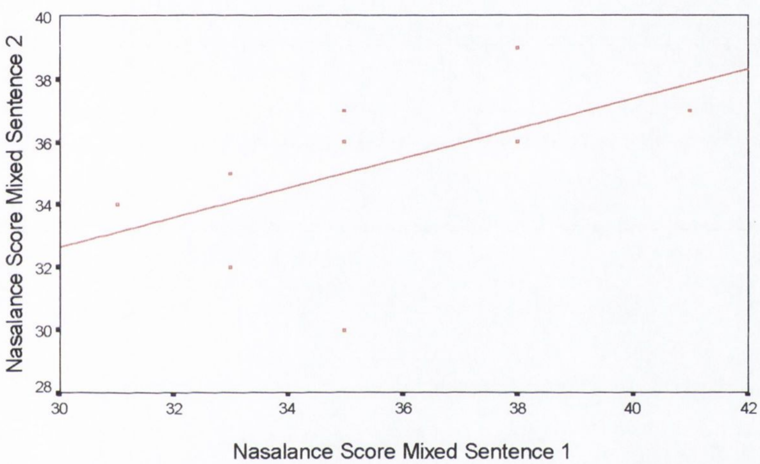


Figure 8d . Test-retest relationship for nasalance scores during the repetition of Mixed Consonant Sentences by 10 normal speakers.

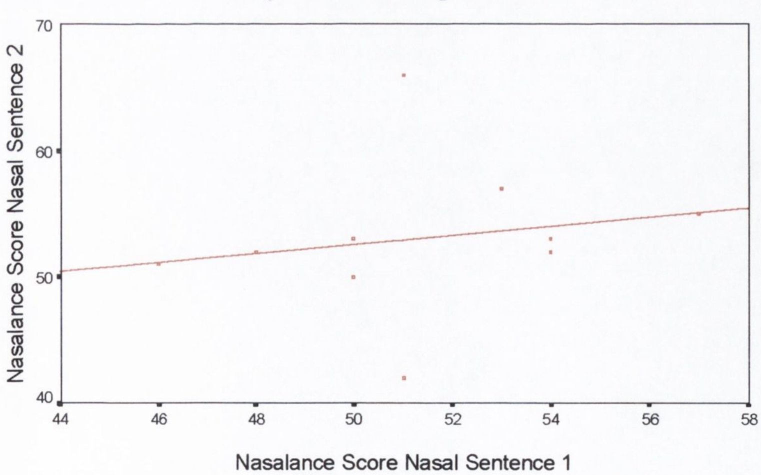


Figure 8e . Test-retest relationship for nasalance scores during the repetition of the Nasal Sentence by 10 normal speakers.

Appendix 9

Analysis of Variance For Nasalance Scores

Source	df	Sums of Squares	Mean Square	F-ratio	Prob
category	4	62950.2	15737.5	1271.8	^{2} 0.0001
gender	1	58.9455	58.9455	0.55045	0.4607
categ*gender	4	76.4244	19.1061	1.5441	0.1897

Appendix 10
Speech Stimulus for Reliability Study

Test Words

Begin, type, fight, cheese, shoot, missing, king, end

Test sentences

1. **Paul likes apple pie**
2. **Ben is a baby boy**
3. **Tim had a tart for tea**
4. **Daddy mended the door** 92
5. **Kevin's looking at the book**
6. **Gary's got a bag of lego**
7. **The phone fell off the shelf**
8. **Vickey's got a very heavy bag**
9. **The shoe shop was shut**
10. **John jumped off the bridge**
11. **The children were watching a football match**
12. **I saw Sam sitting on a bus**
13. **The zebra lives at the zoo**
14. **Mum came home early**
15. **We were away all year**
16. **Will you wear a lily**

Automatic Speech 26

Jack & Jill went up the hill, to fetch a pail of water, Jack fell

down and broke his crown, and Jill came tumbling after.

Counting: 1234567891011121314151617181920
61,62,63,64,65,66,67,68,69,70

Conversational speech

My name is
I live at

I am .....years old

I have .....brothers and ...... sisters.

Appendix 11

Functional Specification for Airflow System

1. Introduction & General Features.

For the purpose of this specification the term nasal airflow refers to the audible flow of air through the nasal cavity during the production of high pressure consonants. Hypernasal resonance refers to the presence of excess nasal resonance on voiced sounds.

This system required should reliably assess nasal airflow during speech sound production and be capable of distinguishing between nasal airflow and hypernasal resonance. The system should also measure oral airflow and similarly be able to distinguish between laminar and turbulent oral airflow.

Systems that facilitate this measurement technique should be P.C. based, reliable, easy to use and of simple construction. The unit is intended to be placed within a speech therapy outpatient clinic and will be used in routine screening paediatric patients pre and post surgery. As such the system should be non invasive and easy to use. The level of co-operation demanded of the patient should be such that it is capable of being used routinely to assess young children (4 - 5 years of age). The system should be capable of being calibrated by a Speech Pathologist.

The measuring system must be accurate, have a linear response and be reproducable so as to provide a reliable quantitative measure of nasal and oral airflow.

The measured data should be capable of being stored on computer disk for archiving. The data from an examination should be capable of being presented graphically. All measured airflows should be plotted against time and the output should include an indication of speech sounds made during the examination as well as identifying the patient and time/date of the examination. The ratio of the nasal to oral airflow should also be capable of being plotted.

The system should be supplied with normative data available for comparison with measured patient data.

The system will also be used as a research tool to correlate objective measures of nasal emission and turbulence with perceptual judgements in terms of degree of airflow, consistency and frequency accross various speech samples.Therefore the system should correlate well with other methods of assessing speech and velopharyngeal function specifically with perceptual judgements of nasal emission and turbulence.

Triona Sweeney

Appendix 12 - Main Pilot Study

Methodology

Participants

The pilot study group consisted of a consecutive series of 10 children who were referred to a national Cleft Palate Unit for investigation of speech problems. Children with cleft palate, velopharyngeal dysfunction, or nasal obstruction, with or without syndromes, were included. All participants presented with abnormal nasality and/or nasal airflow problems, with or without articulatory errors. Age of the participants ranged from 5 to 12;02 years.

Children were excluded if there was evidence of any of the following: severe dyspraxia/dysarthria; learning disability (greater than mild); bilateral hearing loss above 45 db.; nasal blockage associated with an upper respiratory infection; moderate to severe hoarseness of voice; mixed nasal resonance; and an inability to complete the assessment protocol due to behavioural problems or poor co-operation.

Hearing acuity was assessed within one month of the speech assessment.

Materials

A classification system to describe nasality and nasal airflow problems has been developed. The system describes nasality (hypernasality and hyponasality) and nasal airflow in an auditory perceptual framework (Chapter 3, Section 2).

Nasality was assessed on a descriptive scale where the presence or absence of nasality problems were recorded. If present, the deviation from the norm was described. Hypernasality and hyponasality were described as consistent or inconsistent. The presence or absence of cul de sac resonance was recorded.

Nasal airflow errors were divided into 4 categories - nasal emission, nasal fricative, nasal turbulence and velopharyngeal fricative. Each category was described in an auditory perceptual framework where audibility, frequency, consistency and phoneme specificity of the airflow were recorded. (Appendix 13a)

Nasal resonance was measured using the Nasometer model 6200.3 manufactured by Kay Elemetrics. A headset, containing a sound separator with microphones on either side, detected oral and nasal acoustic energy during speech production. The signal was then filtered and digitised and the data was processed by a computer (IBM PC). The resultant signal was a ratio of nasal to nasal-plus-oral acoustic energy and was expressed as a Nasalance score.

Pressure/flow measurements were recorded using the PERCI Speech Aerodynamic Research System (SARS) (Warren, 1994). The system consisted of a nasal flow tube, nasal pressure catheter and oral pressure catheter which connect to pressure and flow transducers. The pressure drop across the velopharyngeal port was measured by placing one catheter in a nostril and another in the mouth. The nasal catheter was secured by a soft foam cork which created a stagnant column of air. Both catheters measured static air pressures and transmitted these pressures to a pressure transducer. Nasal airflow was measured by a heated pneumotachograph connected by plastic tubing.

Procedures

All participants were initially assessed for articulation errors using an adapted phonemic screening test (PACS TOYS, Grunwell and Harding, 1995, NFER-NELSON Publishing Co.) (Methodology, Chapter 5, section 4)

Perceptual Assessment of Nasality and Nasal Airflow Errors

The perceptual assessment was audio recorded using a Sony Scoopman digital audio tape recorder (DAT NT1) with a Sony Electret Condensor Microphone (ECM - 44B). The high quality microphone was placed 4 inches from the child's mouth and the recording volume was set between 24 and 27 so as to obtain sufficient volume without distortion of the speech signal. Volume adjustments were made at the playback stage to obtain optimum output of speech.

The following speech sample was obtained:

1) 16 sentences which the participant repeated after the examiner.

2) A nursery rhyme (Jack and Jill)

3) Counting / repetition of 1 to 20 and 60 to 70.

4) Two minutes of conversation speech.

Each level of speech (word, sentence, automatic speech and conversational speech) was analysed separately from the audio tapes. Before the total speech sample was analysed the judge listened to an Anchor Stimulus. (i.e. a similar shortened speech sample produced by a normal speaker of similar age and the same sex)

Nasality and nasal airflow were described perceptually using the classification system developed for this study (Chapter 4).
Articulatory errors were also recorded.

Nasometry

Nasality was measured using the Nasometer. The headset was placed securely on the child's head ensuring that the sound separator sat under the nose and above the upper lip without interfering with lip movement. The recording rate was set at 4 seconds. There was a one second gap between each sentence in the group and a five second gap between each group to facilitate analysis. Each participant repeated the 16 test sentences individually and each sentence was recorded on the nasometer. Sentences were presented in the following groups:

High pressure consonants;
Low pressure consonants;
Mixed consonants;
Nasal consonants.

Nasalance scores for the total speech sample were calculated. Using a display duration of 20, the initial 5 sentences containing high pressure consonants were marked with the cursors and the nasalance score was calculated. The Low Pressure consonants, the Mixed consonant sentences and the nasal sentence were analysed in the same manner.

Pressure/Flow Measurements

Airflow measurements were made using the PERCI SARS, during the production of *pa pa pa, pi pi pi* and *hamper hamper, hamper.* The system was calibrated at the beginning of each day, and the base level was checked prior to each complete speech assessment. Recordings were made in the *Velopharyngeal Area Measurement* sample design.

The pressure catheter was place in the right nostril. The appropriate size airflow tube was placed in the left nostril. The oral pressure catheter was placed in the child's mouth, ensuring that the tongue did not obstruct the tube. When the tubes were in place, the child was asked to say '*pa pa pa*' 3 times. This speech sample was recorded in the Scope Mode of the programme, which allows the examiner to check that pressure and flow are being recorded accurately. When the graph was consistent, the child repeated the sample again and this was recorded in the Sample mode. Nine recordings of /p/ in *pa* were saved.

This procedure was repeated during the production of '*pi pi pi*' and '*hamper hamper hamper*'. Simultaneous audio recordings of the speech sample were made to validate analysis. All pressure/flow measurements were saved, and later retrieved for analysis.

Analysis was made in the Retrieve Mode of the SARS programme. The speech playback was then selected to help identify clear productions of the oral consonant /p/. The Velopharyngeal Area Measurement was selected and the cursor was placed at the 6 highest peak oral pressure points on the graph. When the 6 points had been selected, the Mean measures of oral pressure, nasal pressure, nasal flow and velopharyngeal port area size for the 6 samples were calculated and printed in Quick Report Mode.

.

Information on Articulation Summary, Voice Quality, Orofacial Assessment, Hearing, Language and Cognitive Development was recorded (Appendix 17).

Results

10 consecutive children who fell into the inclusion criteria were assessed using the above procedures. There were 6 girls and 4 boys with a mean age of 7 years 5 months (ranging from 5 years to 12 years; 2 months). Age range and diagnosis are presented in Table 1.

Table 1. Age, gender and diagnosis of the pilot study group.

Participant	Age	Sex	Classification
1	11	F	submucous cleft palate
2	9.11	M	velopharyngeal insuffic
3	7.11	M	unilat cleft palate
4	12.2	F	secondary cleft palate
5	5.6	M	phoneme specific vpi
6	6.11	F	velopharyngeal insuffic
7	7.6	M	secondary cleft palate
8	5.6	F	velopharyngeal insuffic
9	5	F	secondary cleft palate
10	5.11	F	secondary cleft palate

Perceptual Results

Nasality

Eight out of the ten children were considered to have Hypernasal speech. None were hyponasal.

Table 2. Perceptual ratings of Hypernasality for 10 participants in the pilot study.

Participant	Nasality	Consistency
1	Absent	consistent
2	mild/acceptable	consistent
3	mild/moderate	inconsistent
4	mild/moderate	consistent
5	Absent	consistent
6	mild/moderate	inconsistent
7	moderate/severe	inconsistent
8	moderate	consistent
9	mild/acceptable	inconsistent
10	moderate/severe	inconsistent

Nasal Airflow

Two children were perceived as having nasal emission (Table 3). Eight had nasal turbulence (Table 4). Three had velopharyngeal fricatives (Table 5). None had nasal fricative.

Table 3. Perceptual ratings of Nasal Emission

Participant	Strength	Consistency	Frequency	Phoneme specific
4	weak	inconsistent	infrequent	0
10	weak	consistent	frequent	0

Table 4. Perceptual ratings of Nasal Turbulence

Participant	Strength	Consistency	Frequency	Phoneme Specific
1	weak	consistent	frequent	0
2	weak	consistent	frequent	0
3	weak	consistent	frequent	0
4	weak	consistent	frequent	0
6	weak	consistent	frequent	0
7	weak	inconsistent	infrequent	0
8	weak	consistent	frequent	0
9	weak	consistent	frequent	0
10	weak	consistent	frequent	0

Table 5. Perceptual ratings of Velopharyngeal Fricatives

Participant	Strength	Consistency	Frequency	Phoneme Specific
5	Strong	consistent	frequent	yes
6	Weak	inconsistent	infrequent	yes
8	Weak	consistent	frequent	yes

Nasometry Results

Nasalance scores were calculated for the Total Test sentences and each sentence category (Table 6).

Table 6. Nasalance scores for all participants during the production of each sentence category.

Participants	Total	High Pressure Consonants	Low Pressure Consonants	Mixed Consonants	Nasal Consonants
Normal Mean/ SD	25 / 5	14 / 5	16 / 6	34 / 7	51 / 7
1	35	16	29	43	62
2	49	42	51	51	63
3	30	24	21	34	50
4	62	55	63	64	71
5	28	24	9	35	46
6	51	43	53	53	69
7	53	37	57	59	67
8	63	56	62	67	70
9	66	62	64	68	68
10	64	62	68	64	64

Total sentence sample

Nasalance scores on the total sentence sample ranged from 28% to 66%. All cases had nasalance scores above the normal scores for Irish children (mean nasalance score for Irish children is 25%, ranging from 16% to 35%, Chapter 3). However, participants 3 and 5 had a score of 28% and 30% respectively which is within 1 standard deviation from the norm. Participant 1 had a nasalance score of 35%, two standard deviations from the norm. The mean score for the pilot group is significantly higher than the mean for normal speakers ($p<.001$).

High Pressure consonants

A wide range of nasalance scores was noted within this category (16% to 62%). Participant 1 had a nasalance score of 16% on the high pressure sample, which is within normal limits. The mean score for the pilot group is significantly higher than the mean score for normal speakers ($p<.001$)

Low Pressure consonants

The nasalance scores in this category ranged from 9% to 68%, indicating a substantial variation in scores. The mean score for this group was significantly higher than normal scores and with a greater range ($p =.001$). Participant 5 had a lower nasalance score than the normal mean; however, the score was within the normal range. All participants except participants 3 and 5, had higher nasalance scores for the low pressure sentence category than for the high pressure sentences.

Mixed consonants

Nasalance scores ranged from 34% to 68%. Scores on total and mixed sentence type were similar, as expected. Participants 3 and 5 had scores similar to the normal group. The mean score for the pilot group was significantly higher than the normal mean ($p <.001$).

Nasal consonants

Nasalance scores for the nasal sentence category were all within the normal range.

Pressure/Flow Results

Nasal Flow

Normal speakers have been found to have mean nasal airflow ranging from 1 mls to 4 ml/sec during the production of /p/ in 'pa', 'pi' and 'hamper' (Zajac et al., 1997). In this study, all participants except two (participants 5 and 9) had nasal flows above the normal mean.

Table 8. Mean Nasal Flow Measurements (mls) for ten participants during the production of /p/ in 'pa', 'pi' and 'hamper'

Participants	**'pa'**	**'pi'**	**'mp'**
1	27	25	31
2	57	55	67
3	4	76	15
4	107	123	113
5	1	2	1
6	4	45	4
7	0	0	22
8	5	49	34
9	7	1	0
10	10	37	10

These results indicated variation between speakers, with nasal flows ranging from 0 mls in participant 7 to 123 mls in participant 4. Nasal flow measurements on /p/ in 'pi' were greater than nasal flow measurements in 'pa' for participants 3,4,6,8,10. There was a wide range of nasal flow measurements during the production of the three different speech samples. For example, participant 3 had flow measurements of 4 mls on 'pa', 15 mls on hamper and 76 mls on 'pi'.

Oral Pressure

Oral pressure measurements for normal speakers ranges from 6.2 to 8 cm H_2O during production of /p/ in 'pa' 'pi' and 'hamper' (Zajac et al., 1997). Warren (1979) states that oral pressure above 3 cm H_2O is considered normal. All participants except one had normal oral pressure scores (Table 9). Participant 4 had oral pressure below normal ranging from 1.7 to 2.5 cm H_2O.

Table 9. Mean Oral Pressure measurements (cm H_2O) for ten participants during the production of /p/ in 'pa', 'pi' and 'hamper'

participants	**'pa'**	**'pi'**	**'mp'**
1	5.2	5.8	4.7
2	8	7.9	6.2
3	7.2	4.4	6.6
4	2.5	1.9	1.7
5	8.4	7.9	5.7
6	5.6	7.8	4.3
7	10.5	9.8	8.7
8	5.4	7.9	5.4
9	9.4	5.6	3.7
10	6.1	6.7	6.9

Velopharyngeal Port Area

Velopharyngeal Port Area measurements were calculated by the PERCI SARS software for each participant. However, results must be considered with caution due to unstable baseline graphs for many participants.

Table 10. Velopharyngeal Port Area Measurements (cm^2) for the 10 participants during the production of /p/ in 'pi', 'pa' and 'hamper' measurements.

Participants	'pa'	'pi'	'mp'
1	0.037	0.026	0.03
2	0.04	0.034	0.04
3	0.058	0.306	0.092
4	0.118	0.262	0.161
5	0.004	0.004	0.005
6	0.004	0.02	0.003
7	0.001	0.008	0.015
8	0.016	0.085	0.032
9	0.01	0.018	0.018
10	0.015	0.017	0.004

The Relationship Between Measurements

Correlations were calculated to assess the relationship between perceptual measures of nasality and nasalance scores and between perceptual measures of nasal airflow and pressure/flow measures.

Nasality and Nasalance Scores

The Spearman Rank correlations were calculated to assess whether there was an association between perceived hypernasality ratings and nasalance scores (Table 10). It must be remembered that this correlation is based on a small number of participants. A correlation of 0.53 indicated a moderate, positive correlation between perceptual ratings of nasality and nasalance scores on the Total Test sentences.

Table 10. Spearman correlation coefficients for nasality ratings and nasalance scores for each sentence category.

Nasality	Total nasalance	High pressure nasalance	Low pressure nasalance	Mixed nasalance
correlation	.53	.41	.58	.41

Correlations were all positive and ranged from fair (r =.41) to moderate (.58). None were significant. The strongest correlation was between perceptual ratings of nasality and nasalance scores on low pressure consonant sentences.

Nasal airflow and Pressure/Flow Measures

For the purpose of analysis perceptual ratings of nasal airflow were divided into the four main groups - nasal emission, nasal fricative, nasal turbulence and velopharyngeal fricative. Each group was given a score from 1 to 8 according to descriptions of strength, frequency, and consistency (Methodology, Chapter 5).

Nasal airflow was compared to nasal flow measures on all three speech samples (‘pa’, pi’ ‘mp’) using Spearman rank correlations (Table 11)

Table 11. Correlation coefficients for perceptual ratings of nasal airflow errors and nasal flow measurements during the production of /p/ in ‘pa’, ‘pi’ and ‘hamper’

Nasal airflow	**‘pa’**	**‘pi’**	**‘mp’**
nasal emission	r = .49	r = .30	r = .01
nasal turbulence	r = .68	r = .68	r = .28
velopharyngeal fricative	r = -.43	r = - .17	r = -.27

Results indicate that correlation coefficients ranged from a moderate positive correlation (r = .68) to a poor negative correlation (r = -.43). The strongest correlation was found between nasal turbulence and nasal flow measures on ‘pa’ and ‘pi’.

Weak correlations were found between perceptual ratings of nasal emission and oral pressures on all speech samples (Table 12). A moderate positive correlation was found between nasal turbulence and oral pressure on ‘pi’. None of the correlations were statistically significant.

Table 12. Correlation coefficients for perceptual ratings of nasal airflow errors and oral pressure measurements during the production of /p/ in 'pa', 'pi' and 'hamper'

Oral Pressure	'pa'	'pi'	'mp'
nasal emission	r = -.39	r = - .40	r = -.01
nasal turbulence	r = -.39	r = -.59	r = -.58
velopharyngeal fricative	r = -.03	r = .41	r = -.17

Correlation coefficients were not calculated for velopharyngeal port area or for differential pressure measurements due to unreliable pressure/flow results.

Methodological problems and solutions

A number of methodological problems became apparent during the pilot study. In the perceptual section of the protocol, nasal turbulence was noted when listening to speech on the audio recorder and headphones, where it had not been detected during live assessment. It may be that the nasal turbulence was too weak to be noted in a live analysis but could be detected if the signal was amplified. If nasal turbulence was detected on the audio recorder, but not in the live situation, the recording level was lowered to reflect the live speech quality.

For the pressure/flow assessment the airflow tube was initially placed in the left nostril and the pressure catheter was placed in the right nostril. If there was some nasal blockage due to a deviated septum, the flow graph had less peaks than expected. In one case it was impossible to analyse the data.

Eighty five percent of the graphs were not at 0 baseline during the assessment, despite prior calibration. This was not evident during data collection, but became apparent during analysis. Corrections were made by measuring the baseline and calculating the appropriate pressures and flows. However accurate calculation of the velopharyngeal port area and differential pressure was not possible. One possible explanation for the change in baseline may be that many of the participants inhaled and/or exhaled through their noses while the system was being prepared for data sampling. The respiratory airflow may have changed the baseline from 0.

Discussion

Due to the small number of participants in the pilot study any initial results must be considered with caution.

Nasality/NasalAirflow

Perceptual ratings of nasality ranged from normal resonance to moderate/severe hypernasality. Hypernasality was inconsistent in fifty percent of participants. Only two participants were found to have nasal emission. Six out of eight participants, who were reported to have nasal turbulence, were described as having weak, consistent and frequent nasal turbulence. In the velopharyngeal fricative group the strength, consistency and frequency varied.

Nasalance

All participants had increased nasalance scores above the normal mean in one or more of the sentence categories. Although participants 1 and 5 were reported to have normal resonance, each had an increased nasalance score on one category. Participant 5 had a nasalance score of 24% on the high pressure consonant sentences (mean for normal speakers is 14%). This may be a result of his phoneme specific velopharyngeal fricative which may increase the score on the high pressure sample, as the nasometer may not distinguish between turbulent sound and hypernasality (Karnell, 1995).

Participants 1,3 and 5 had nasalance scores within the norm for the total speech sample; however, within sentence categories higher nasalance scores were noted. These higher scores may reflect inconsistencies in speech or additional airflow problems associated with velopharyngeal dysfunction.

In analysing the nasalance scores one would expect the group with nasal turbulence to have higher nasalance scores for high pressure consonants than for low pressure consonants. Karnell (1995) studied 43 patients and found important differences in mean nasalance measurements for some participants when high pressure sentences and low pressure sentences were compared (Chapter 2, Section 3.4 (i)). He hypothesised that participants with nasal turbulence/nasal emission would have higher nasalancce scores on high pressure consonant sentences compared to low pressure consonant sentences. This was not the case in the pilot study. One explanation for the difference between these results and those in Karnell's study may be that all participants in this study had mild nasal turbulence, which may not have been detected by the nasometer.

Nasal Airflow

Eight out of the ten participants had nasal flow measurements above the normal mean. Participant 5 was perceived to have a phoneme specific velopharyngeal fricative on /s/ and /z/, and as expected had no nasal flow during the production of /p/. Participant 9, who also had normal nasal flow on /p/, was perceived to have weak, consistent and frequent nasal turbulence during speech. It is possible that the weak nasal turbulence was not present during the production of the limited speech sample used for pressure/flow measurements, but was present on the more detailed speech samples used for perceptual analysis. This highlights one of the main problems in comparing two assessments that include different speech samples.

Participants 4 and 10 had nasal emission and nasal turbulence. Participant 4 had the highest nasal flow of the group (107 mls to 123mls) and also had mild/moderate hypernasality. However, participant 10 had moderate to severe hypernasality with nasal emission and turbulence but had minimal increase in nasal flow (10mls to 37mls).

Nasal flow measures varied across three speech samples in four participants (Table 13). Variations in nasal flow measures in these four participants may well be explained by the perceptual results. Each participant had an inconsistent speech pattern. Participant 3 had inconsistent hypernasality and inconsistent nasal turbulence, participant 6 had inconsistent nasal turbulence and inconsistent velopharyngeal fricative, participant 8 had inconsistent nasal turbulence and participant 10 had inconsistent nasal emission and turbulence.

Table 13. Variations in nasal flow measures across three speech samples.

Participant	Pa	Pi	Hamper
3	4	76	15
6	4	45	4
8	5	49	34
10	10	37	10

One would have expected the greatest nasal flow measurements on the production of the nasal plosive cluster in hamper, as the speaker has to shift rapidly from a nasal to an oral production (Warren, 1979). However the greatest nasal flow measurements were noted on 'pi'. This may be explained by the fact that velopharyngeal incompetence has been found to be more marked on high vowels in cases with borderline velopharyngeal insufficiency due to the downward pull of the palate by the palatoglossus during high vowel production. (Smith & Guyette, 1996)

Correlation between nasality and nasalance scores

Initial results indicated a moderate positive correlation between nasality and nasalance scores however correlations were not statistically significant. No conclusions can be made from the statistical correlation due to the small number of cases in the pilot study. Two participants were found to have normal nasality on perceptual assessment. One of these participants had a nasalance score within the norm. The other participant had a nasalance score above the normal mean. However, she had weak, inconsistent and infrequent nasal turbulence, which may have elevated the nasalance score (Karnell, 1995). One participant had mild/moderate hypernasality on the perceptual classification and a nasalance score of 35%, which was expected. Three participants had moderate or moderate/severe hypernasality and high nasalance scores as expected. Four participants fell into the normal/ mild categories but had high nasalance scores ranging from 50% to 65%.

Correlation between pressure/flow measures and perceptual judgements of nasal airflow.

Correlations between nasal airflow categories and nasal flow measurements ranged from moderate to poor. Again, no conclusion can be made due to the small number of children tested in the study. Poor correlations may partly be due to methodological problems in placing the tubes and obtaining a steady baseline on computer.

Warren (1998, personal communication) reported that participants with borderline velopharyngeal function were notoriously difficult to assess objectively. In the pilot study many of the participants had weak nasal turbulence, with varying degrees of hypernasality. The fact that the nasal airflow errors were mild may indicate borderline velopharyngeal function for airflow. Initial results indicate a moderate positive correlation between nasal turbulence and nasal emission and measurements of nasal flow. Although all participants with nasal turbulence had some nasal flow, the range of flow varied from 2.8 to 123 ml/sec. Variations in nasal flow may be explained by incorrect placement of the tube in some cases where there was nasal blockage.

The aim of this pilot study was to evaluate the procedures for the main study. Some changes in methodology are recommended for the main study.

Recommendations for Main Study

As a result of the intra-rater and inter-rater studies (Chapter 4), two changes in the procedures for perceptual assessment are recommended. During the perceptual assessment the anchor stimuli will not be used and ratings will be made during the live assessment sessions. Audio recordings will be made at the same time for detailed analysis if required.

For pressure/flow measurements, each participant in the main study will be assessed to ascertain which nostril is least blocked for appropriate placement of the nasal flow tube. (Chapter 5, Methodology).

In order to ensure a stable baseline in pressure/flow measurements, the system will be calibrated prior to each assessment session. Before each participant is assessed the system will be checked to evaluate the pressure and flow baselines. If the baseline is not at 0, the system will be calibrated again.

Appendix 13a Complete Assessment Sheet

Test Words

Pa Pa Pa, Pi Pi Pi, Hamper Hamper Hamper

Test sentences

1. **Paul likes apple pie**
2. **Ben is a baby boy**
3. **Tim had a tart for tea**
4. **Daddy mended the door** 92
5. **Kevin's looking at the book**
6. **Gary's got a bag of lego**
7. **The phone fell off the shelf**
8. **Vickey's got a very heavy bag**
9. **The shoe shop was shut**
10. **John jumped off the bridge**
11. **The children were watching a football match**
12 **I saw Sam sitting on a bus**
13. **The zebra lives at the zoo**
14. **Mum came home early**
15. **We were away all year**
16. **Will you wear a lily**

Automatic Speech 26

Jack & Jill went up the hill, to fetch a pail of water, Jack fell

down and broke his crown, and Jill came tumbling after.

Counting: **1234567891011121314151617181920**
61,62,63,64,65,66,67,68,69,70

Conversational speech

My name is
I live at

I am .....years old

I have .....brothers and ...... sisters.

Appendix 13b

NASALITY / NASAL AIRFLOW ASSESSMENT

Nasality

Hypernasality: Present Absent
a) mild, evident but acceptable.
b) mild/moderate, unacceptable distortion , evident on high vowels
c) moderate, evident on high and low vowels.
d) moderate / severe, evident on all vowels and some consonants.
e) severe, evident on all vowels and most voiced consonants.
Consistent Inconsistent

Hyponasality: Present Absent
a) evident, but acceptable
b) moderate - all vowels reduced nasality.
c) severe - total denasal production of nasal consonants.
Consistent Inconsistent

Cul de Sac: Present Absent

Nasal Airflow

Nasal Emission:	Weak	Strong
	Frequent	Infrequent
	Consistent	Inconsistent
	Phoneme Specific	
Nasal Fricative:	Weak	Strong
	Frequent	Infrequent
	Consistent	Inconsistent
	Phoneme Specific	
Nasal Turbulence:	Weak	Strong
	Frequent	Infrequent
	Consistent	Inconsistent
	Phoneme Specific	
Velopharyngeal Fricative:	Weak	Strong
	Frequent	Infrequent
	Consistent	Inconsistent
	Phoneme Specific	
Intranasal turbulence	Present	Absent

Weak / Strong - intensity of audible nasal airflow on sounds.
Consistency - same level and frequency in different speech situations.
Frequency - percentage of nasal emissions in a passage.
Phoneme specific - only present on certain sounds consistently.

Voice: Normal Dysphonic Reduced Volume

Test Words
Articulation ____________________

Nasality ____________________

Sentence Repetition
Articulation ____________________

Nasality ____________________

Automatic Speech
Jack / Jill ____________________

Counting 1 - 20 __________ 60 -70 __________

Conversational Speech
Articulation ____________________

Nasality ____________________

Appendix 14

Calibration of Equipment

Calibration of the Nasometer

To perform the calibration, the head set was inserted into the calibration stand. The calibration stand was placed 12 inches from the left side panel of the Nasometer and in direct line with the calibration speaker (Figure 1). Using the Nasometer programme, 'System' was selected on the main menu. 'Calibrated Headset' was then selected. The instrument was calibrated by adjusting the adjustment screw on the side panel of the Nasometer. The Bar Mode Display was selected. When the line on the display window of the computer was at 50% nasalance, calibration was completed by pressing the ENTER key.

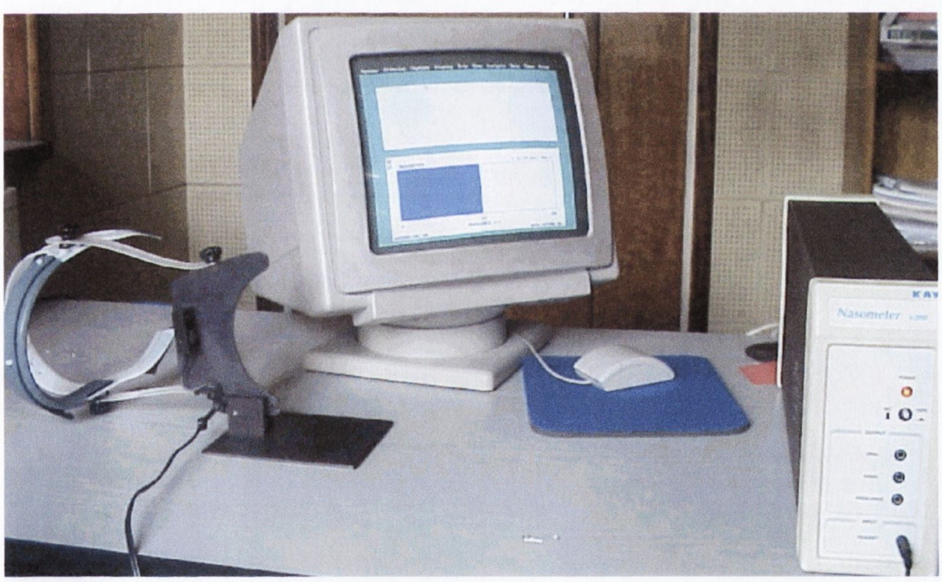


Figure 1 Photograph of calibration set up for Nasometer, indicating the bar mode on the computer screen.

Calibration of PERCI SARS

Calibration was performed by selecting 'Calibration' from the Perci SARS main window. For pressure calibration, 'Pressure' was selected from the drop menu. Initially a zero baseline was established so that the computer program was able to identify what output value from the transducers represented zero. With nothing connected to the pressure transducers, the 'Set' command buttons for pressure 1 and for pressure 2 were selected. The 'Pressure Value' was set at 8.0 for both channels (Figure 2). A syringe was attached to a T tube which connected with the pressure transducers and with a U tube water manometer (Figure 3). The syringe applied 8 cm H_2O pressure which was measured by displacing the water in the U tube manometer by 8 centimetres. The 'Dual' command button was then selected on the Pressure Calibration window. The resulting calibration 'Scale Factor' was compared with the previous calibration. If there was less than 5% variation from the previous calibration, the results were considered acceptable and the 'Save' button was selected. If there was greater than 5% variation between the present calibration and the last calibration, the graphs were checked to ensure that the baseline was at 0. Calibration was then repeated and saved.

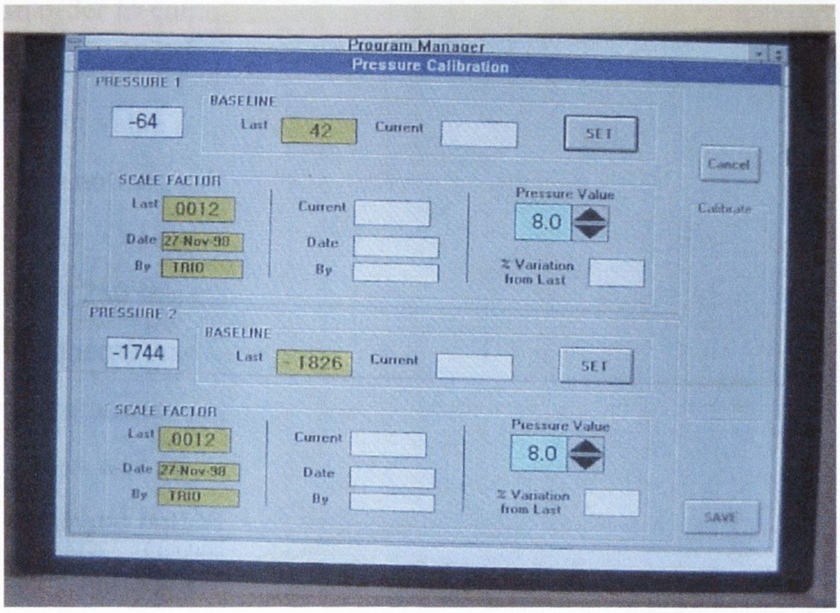


Figure 2. Pressure calibration window of PERCI SARS system.

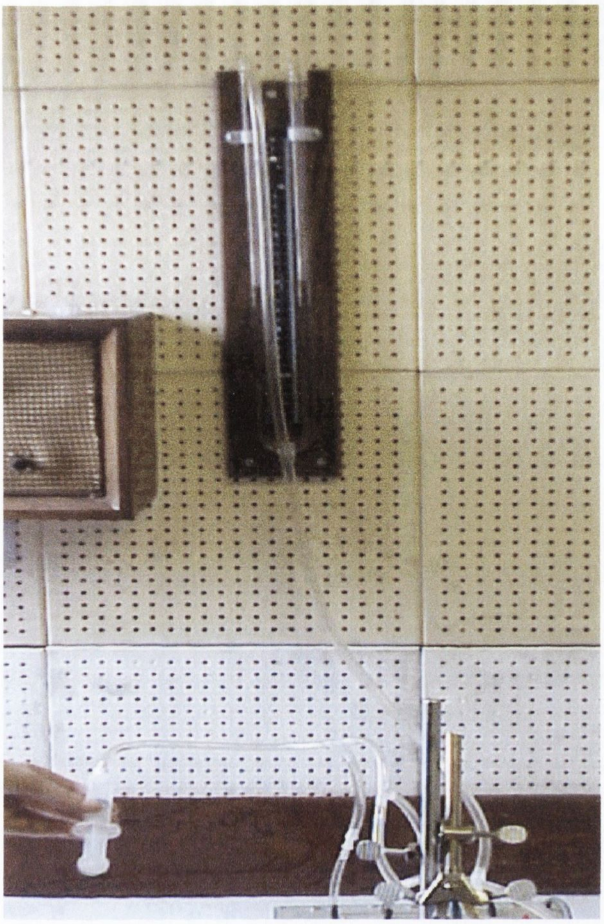


Figure 3. Photograph of the Pressure calibration system for PERCI SARS indicating the U tube manometer, syringe attached to the pressure transducers.

In order to calibrate airflow, the 'Flow' option was selected from the 'Calibration' menu. The zero baseline was established so that the computer program could identify what output value from the pneumotach and flow transducers represented the zero condition. With nothing attached to the pneumotach the 'Set' button for Flow 1 was selected. Under 'Cal methods' the 'Fixed Level' calibration was selected. The Flow Value 1 was set at 250 mls (Figure 4). A tube was connected from the air source to the rotameter and then from the rotameter to the pneumotach. The air source was then turned on until the rotameter reached a reading of 250 mls (Figure 5). The 'Flow Value 1' command button was selected and the air source was turned off. The 'Compute Scale Factor' command button for the Flow 1 channel was selected. If there was less than 5% difference between the present and the previous calibration the 'Save' button was selected. If there was greater than 5% difference between the two calibrations, the procedure was carried out again and the calibration was then saved.

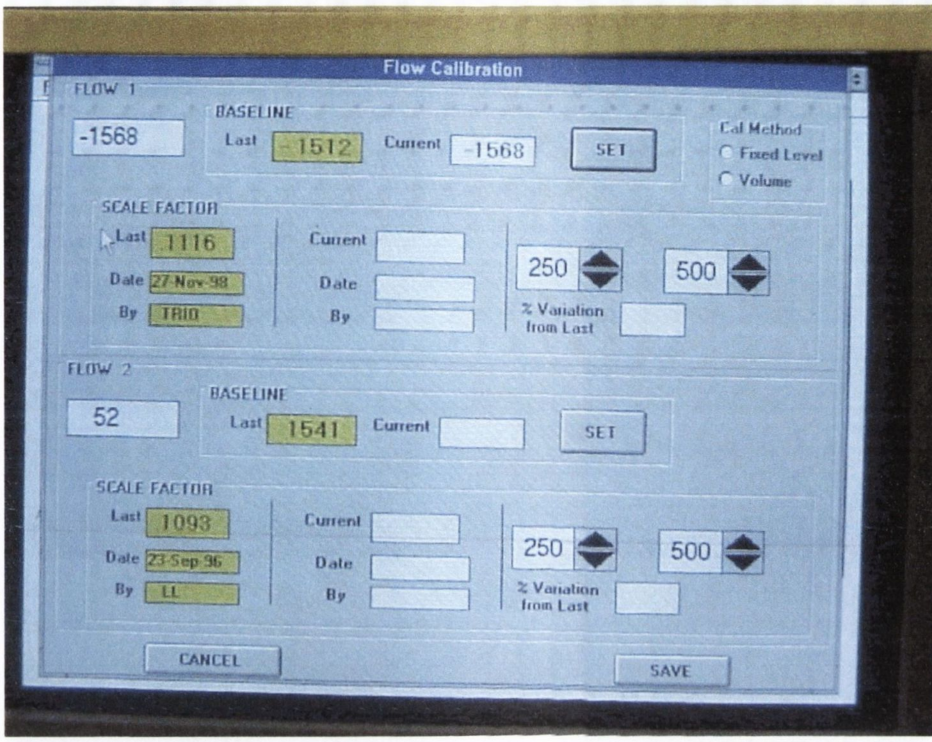


Figure 4. Airflow calibration window of the PERCI SARS system.

Figure 5. Photograph of the Airflow calibration set up for PERCI SARS, indicating the airtank, meters, rotameter and tube attachment for pneumotachograph.

PACS TOYS
SCREENING ASSESSMENT

Word List A – Alphabetical

NFER NELSON

Name .... Chronological Age .... Date .... Tester ....

apple	fork	sand
badge	glove	scissors
brush	go	sheep
bucket	hand	shoe
bus	horse	snake
car	jam/jar	soap
driver	knife	sock
caravan	letter	spade
case	stamp	thumb
chair	nose	tiger
cheese	picture	toe
dinosaur	house	torch
doll	door	trousers
girl	roof	van
mouth	pig	washing machine
elephant	plaster	watch
feather	purse	whistle
fire engine	money	yes
ladder	rabbit	zip
flower	ring	

Additional target words
boat
digger
Mickey Mouse
plane
straw
Thomas
Tank Engine

Additional transcription ______________________

First Published 1995. © Pamela Grunwell and Anne Harding, 1995
Published by The NFER-NELSON Publishing Company Ltd, Darville House, 2 Oxford Road East, Windsor, Berkshire SL4 1DF, UK.
Code 4035 036 (Manual and Cards)/4035 026 (Starter Set) 1(6.95)

Appendix 16

PACS TOYS
SCREENING ASSESSMENT

Extended Phoneme Realization Chart

Name .... Chronological Age .... Date .... Tester ....

	SIWI					SIWW					SFWF				
	✔	Incorrect				✔	Incorrect				✔	Incorrect			
Target Phoneme	Correct	Almost Mature	Develop-mentally Immature	Atypical	Other	Correct	Almost Mature	Develop-mentally Immature	Atypical	Other	Correct	Almost Mature	Develop-mentally Immature	Atypical	Other
m															
n															
ŋ															
p															
b															
t															
d															
ʧ															
ʤ															
k															
g															
f															
v															
θ															
ð															
s															
z															
ʃ															
ʒ															
w															
r															
l															
j															
h															

Published by The NFER-NELSON Publishing Company Ltd, Darville House, 2 Oxford Road East, Windsor, Berkshire SL4 1DF, UK.
Code 4035 036 (Manual and Cards)/4035 026 (Starter Set) 1(6.95)

Appendix 17 Assessment Summary Form

Name **Age** **D.O.B.** **DOA**

Diagnosis **Age of Surgery**

Nasality

Hypernasality: Present Absent
a) mild, evident but acceptable.
b) mild/moderate, unacceptable distortion , evident on high vowels
c) moderate, evident on high and low vowels.
d) moderate / severe, evident on all vowels and some consonants.
e) severe, evident on all vowels and most voiced consonants.

Consistent Inconsistent

Hyponasality: Present Absent
a) evident, but acceptable
b) moderate - all vowels reduced nasality.
c) severe - total denasal production of nasal consonants.

Consistent Inconsistent

Cul de Sac: Present Absent

Nasal Airflow

Nasal Emission: Weak / Strong; Frequent / Infrequent; Consistent / Inconsistent; Phoneme Specific

Nasal Fricative: Weak / Strong; Frequent / Infrequent; Consistent / Inconsistent; Phoneme Specific

Nasal Turbulence: Weak / Strong; Frequent / Infrequent; Consistent / Inconsistent; Phoneme Specific

Velopharyngeal Fricative: Weak / Strong; Frequent / Infrequent; Consistent / Inconsistent; Phoneme Specific

Intranasal Turbulence: Present Absent

---

	pa				pi				mp			
Airflow	OP	DP	NF	VPA	OP	DP	NF	VPA	OP	NP	NF	VPA

Nasalance **Total** **HP** **LP** **Mixed** **Nasal**

Appendix 17a

Articulation (Adapted from GOS.SP.ASS, Sell, Harding & Grunwell, 1998)

CLEFT TYPE SPEECH CHARACTERISTICS:

0	Dentalisation......................	6	Glottal articulation...............
1	Lateralization......................	8	Weak/nasalization...............
2	Palatalization......................	10	Absent pressure con............
2/3	Double articulation...............	11	Gliding of fricative/affric........
3	Backing to velar..................		
4	Backing to uvular................		
5	Pharyngeal articulation..........		

IMITATION: m p b f v n l t d s z . . . n k g h

DEVELOPMENTAL ERRORS ....

Voice	Normal	Dysphonic	Reduced Volume

Orofacial (Adapted from GOS. SPASS, Sell, Harding & Grunwell 1997)

1. nose:	NAD	deviated septum	obstructed	alar abn
2. lips:	NAD	restricted movement	open posture	
3. occlusion:	I	II	III	open bite
4. dentition:	NAD	supernumeracy	missing teeth	malaligned
5. tongue:	mobility	posture	humping	tie
6. palatal fistula:	present	absent		
7. fistula size:	1 < 2mm	2 bet. 2-5mm	3 bet 5-8mm	
		4 > 8mm	5 complete breakdown	

7. palatal structure: ....

8. palatal mobility: ....

9. nasopharynx - tonsils ....
deep pharynx ....
pharyngeal wall movement ....
pharygnoplasty ....

Hearing	normal	loss	
Language Development	normal	delayed	disordered
Cognitive Development	normal	delayed	

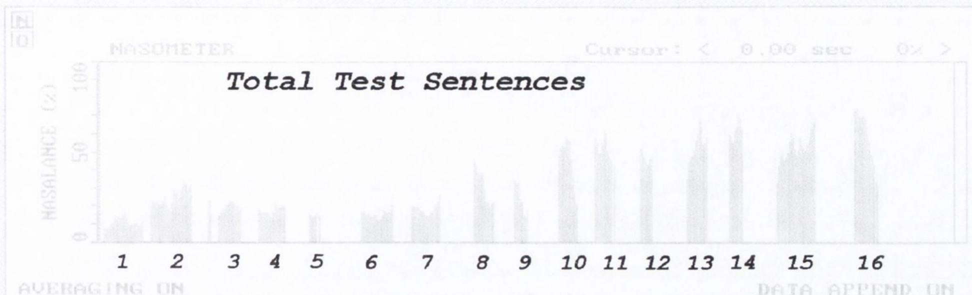
RECOMENDATIONS:

Appendix 18

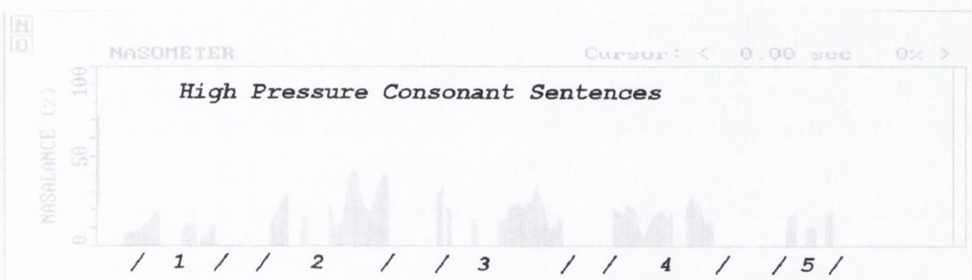
Nasometry graphs indicating nasalance for normal speakers are presented in appendices 18a and 18b. Note that the graph for the Total Test sentences has a duration of 40 seconds, the High Pressure consonant and Low Pressure consonant sentence graphs have durations of 12 seconds and the Nasal sentence has a duration of 4 seconds. Sentences are numbered on each graph and the nasalance scores for each category are presented.

Appendix 18a

Nasometry Graph of Total Test Sentences and High Pressure Consonant Sentences for a normal speaker



Duration: 40 (sec.) Nasalance Score: 31%



Duration: 12 (sec.) Nasalance Score: 19%

<u>Total Test sentences</u>

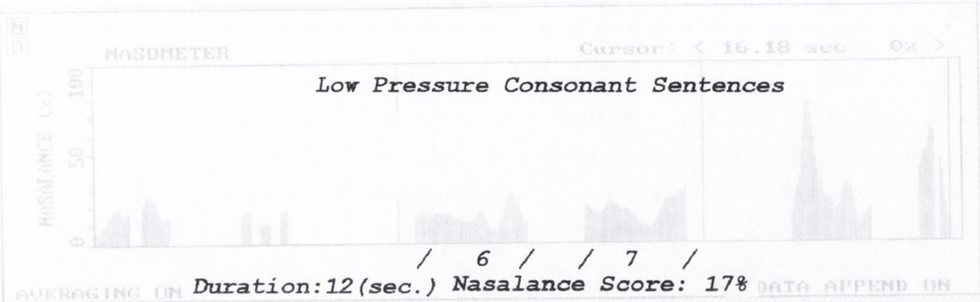
1. Paul likes apple pie
2. Gary's got a bag of lego
3. Vickey's got a very heavy bag
4. The zebra lives at the zoo
5. The shoe shop was shut
6. We were away all year
7. Will you wear a lily?
8. Ben is a baby boy
9. Tim had a tart for tea
10. Daddy mended the door
11. Kevin's looking at the book
12. The phone fell off the shelf
13. I saw Sam sitting on a bus
14. John jumped off the bridge
15. The children were watching a football match
16. Mum came home early

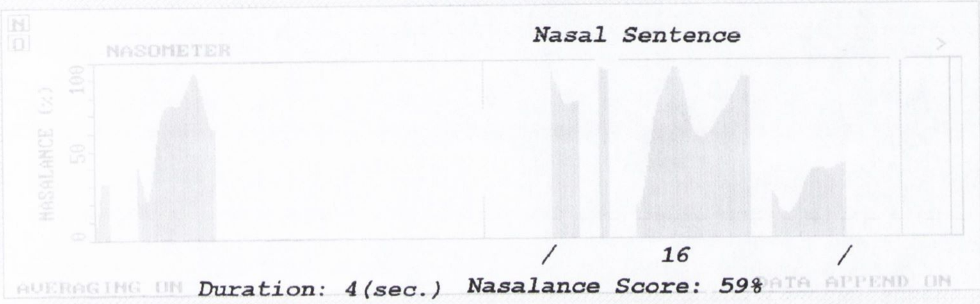
<u>High Pressure Consonant Sentences</u>

Sentence 1 to 5

Appendix 18 b

Nasometry Graph of Low Pressure Consonant Sentences and a Nasal Sentence for a normal speaker





<u>Low Pressure Consonant Sentences</u>

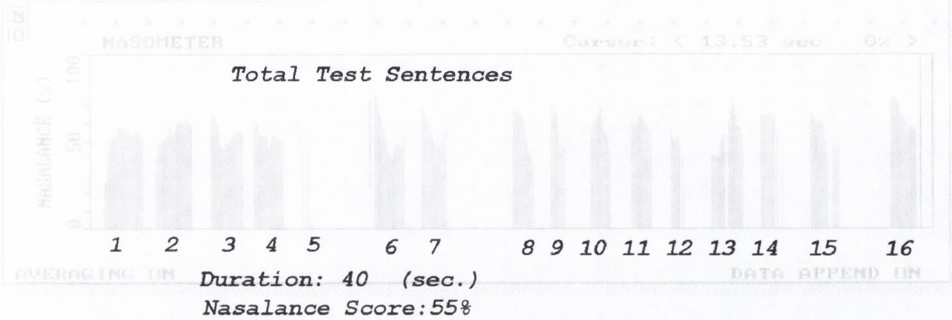
6. We were away all year
7. Will you wear a lily?

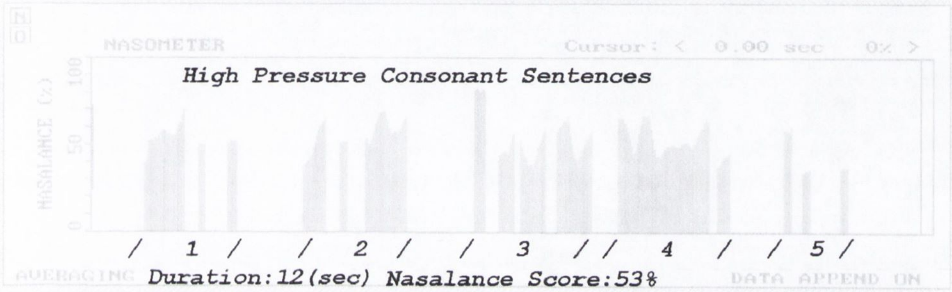
<u>Nasal Sentence</u>

16. Mum came home early

Appendix 18c

Nasometry Graph of Total Test Sentences and High Pressure Consonant Sentences for a hypernasal speaker

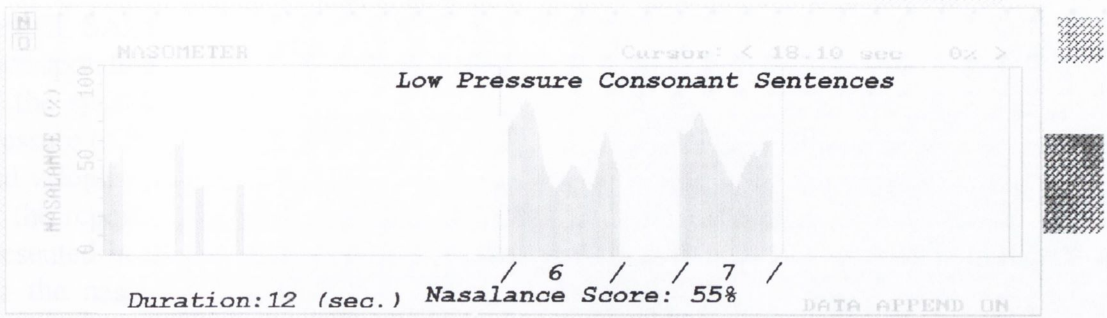




Total Test sentences		High Pressure Consonant Sentences
1.	Paul likes apple pie	Sentence 1 to 5
2.	Gary's got a bag of lego	
3.	Vickey's got a very heavy bag	
4.	The zebra lives at the zoo	
5.	The shoe shop was shut	
6.	We were away all year	
7.	Will you wear a lily?	
8.	Ben is a baby boy	
9.	Tim had a tart for tea	
10.	Daddy mended the door	
11.	Kevin's looking at the book	
12.	The phone fell off the shelf	
13.	I saw Sam sitting on a bus	
14.	John jumped off the bridge	
15.	The children were watching a football match	
16.	Mum came home early	

Appendix 18d

Nasometry Graph of Low Pressure Consonant Sentences for a hypernasal speaker



Low Pressure Consonant Sentences

6. We were away all year
7. Will you wear a lily?

Appendix 19

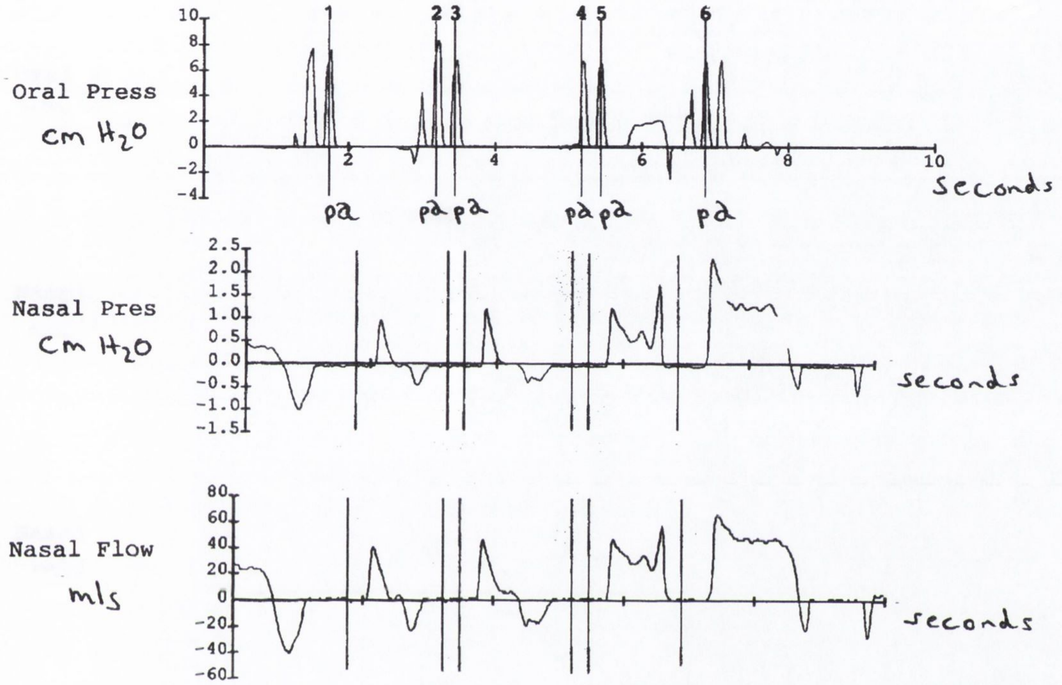
PERCI SARS PRESSURE/FLOW GRAPHS

PERCI SARS graphs and reports with 6 cursors indicating peak oral pressure and corresponding nasal pressure and nasal flow for a normal speaker, during the production of the syllable 'pa' are presented below. Measurements for oral pressure (OP), nasal pressure (NP), nasal flow (Nflow), sound pressure level (SLP), differential pressure (DP) and velopharyngeal port area (AREA) are presented for each cursor in the middle section of the report. The mean measurements, with standard deviation and standard error are presented in the bottom section of the report. Note: there is a time shift of 0.5 seconds for the nasal pressure and nasal flow graphs due to the graphic software used; the cursors do not always correspond to the oral peak in the print out; however, peak points were verified on the computer screen in a yellow measurement box beside the oral pressure graph (see Figure 5.6). Three pressure/flow graphs for a normal speaker during the production of 'pa', 'pi' and 'hamper' are presented in appendices 19a, 19b, 19c. Pressure/flow graphs for a speaker with nasal emission are presented in appendices 19d, 19e and 19f.

Appendix 19a

Pressure/flow graph of 'pa' for a normal speaker

```
'UBLIN CHILDREN'S HOSPITAL   SARS REPORT        DATE:01-09-96    By:TRIO
 Measurement Report From File: K     HP.R01     Type: HAMPER WITH VOICE
---------------------------------------------------------------------------
Patient:                         Hospital ID:                      Visit:
DX:  0         Age:  0Y    0M    Eval Date: 09/01/96          Examiner: TRIO
---------------------------------------------------------------------------
```



Cur# ---------------------- VP Area ------------------------

Cur#							
1	OP=07.6	NP=-00.1	NFlow=000.0	SPL=061.0	DP=07.63	AREA= 0.000	X=1.7
2	OP=08.3	NP=-00.1	NFlow=001.8	SPL=060.0	DP=08.37	AREA= 0.001	X=3.2
3	OP=06.8	NP=-00.1	NFlow=000.0	SPL=060.0	DP=06.80	AREA= 0.000	X=3.4
4	OP=06.7	NP=-00.1	NFlow=000.0	SPL=061.0	DP=06.77	AREA= 0.000	X=5.2
5	OP=06.5	NP=-00.1	NFlow=001.8	SPL=061.0	DP=06.50	AREA= 0.001	X=5.4
6	OP=06.6	NP=-00.1	NFlow=000.0	SPL=061.0	DP=06.63	AREA= 0.000	X=6.8

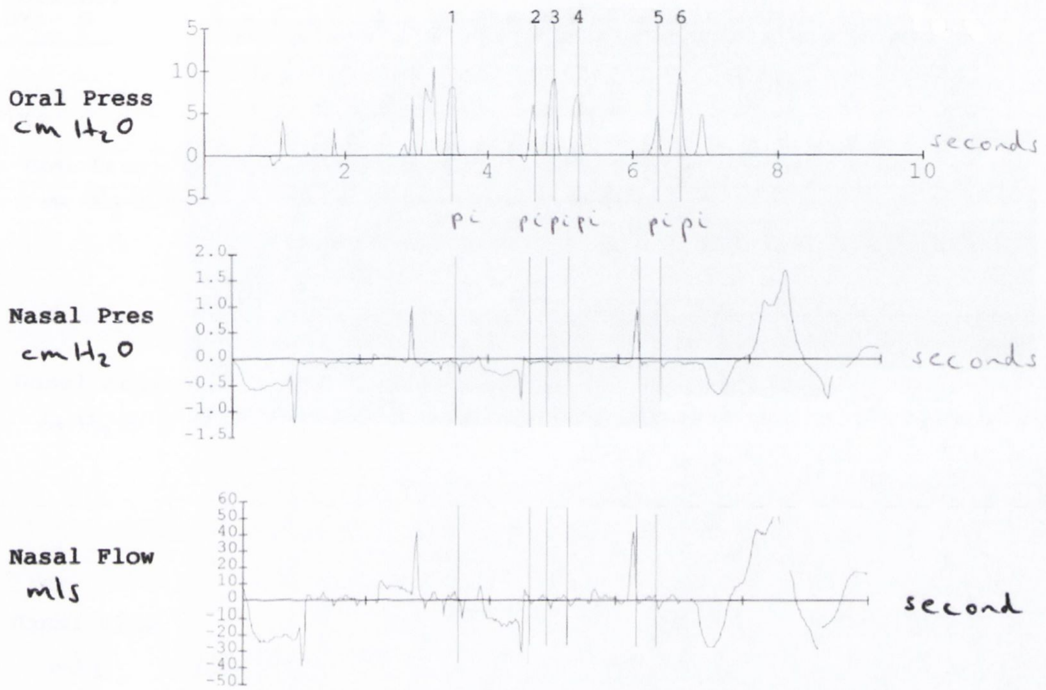
================ VP Area

Oral Press	Mean = 7.1	STD = 00.73	SE= 00.30
Nasal Pres	Mean = -0.1	STD = 00.01	SE= 00.00
Nasal Flow	Mean = 0.6	STD = 00.92	SE= 00.37
SPL	Mean = 60.7	STD = 00.52	SE= 00.21
Computed DP	Mean = 7.12	STD = 00.73	SE= 00.30
AREA(sq cm)	Mean = 0.000	STD = 0.000	SE= 0.000

Appendix 19b

Pressure/flow of 'pi' for a normal speaker

```
   DUBLIN CHILDREN'S HOSPITAL  SARS REPORT       DATE:01-09-96    By:TRIO
  Measurement Report From File: EK    HP.R02      Type: HAMPER WITH VOICE
-------------------------------------------------------------------------------
 Patient:                             Hospital ID:                     Visit:
 DX:  0        Age:  0Y   0M        Eval Date: 09/01/96          Examiner: TRIO
-------------------------------------------------------------------------------
```



Cur#			VP Area				
1	OP=08.0	NP=-00.1	NFlow=001.5	SPL=061.0	DP=08.11	AREA= 0.001	X=3.4
2	OP=06.0	NP=-00.1	NFlow=-000.2	SPL=060.0	DP=06.03	AREA= 0.000	
3	OP=09.1	NP=-00.1	NFlow=001.5	SPL=061.0	DP=09.11	AREA= 0.001	X=4.9
4	OP=07.1	NP=-00.1	NFlow=001.5	SPL=061.0	DP=07.14	AREA= 0.001	X=5.2
5	OP=03.7	NP=00.0	NFlow=008.5	SPL=061.0	DP=03.66	AREA= 0.005	X=6.33
6	OP=09.7	NP=-00.1	NFlow=001.5	SPL=061.0	DP=09.80	AREA= 0.001	X=6.6

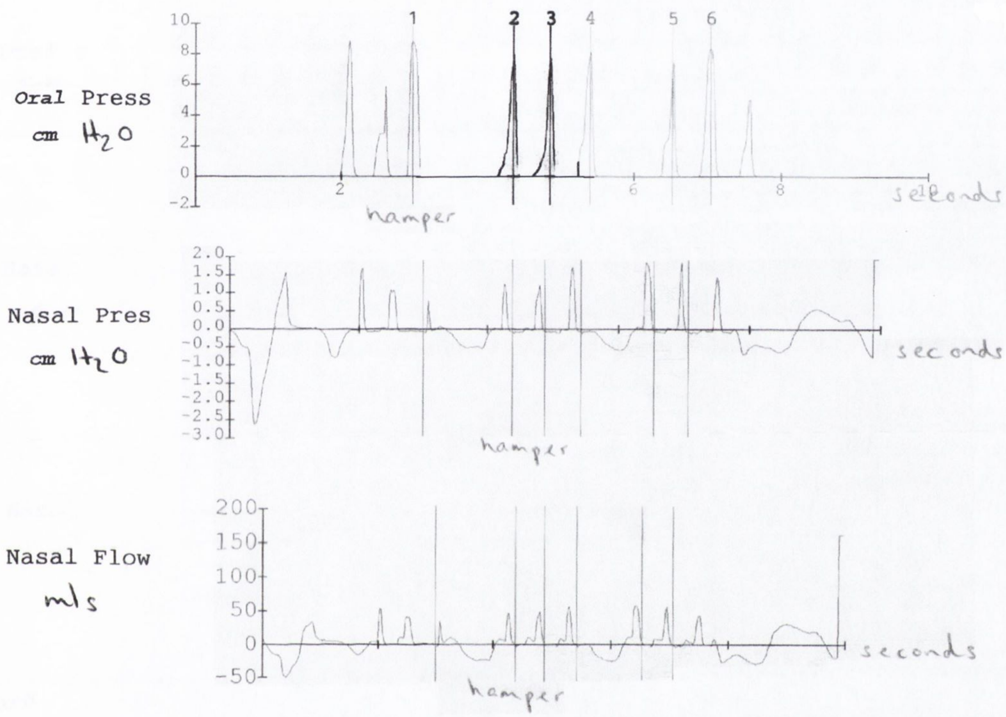
VP Area

Oral Press	Mean = 7.3	STD = 02.21	SE= 00.90
Nasal Pres	Mean = 0.0	STD = 00.04	SE= 00.02
Nasal Flow	Mean = 2.4	STD = 03.07	SE= 01.25
SPL	Mean = 60.8	STD = 00.41	SE= 00.17
Computed DP	Mean = 7.31	STD = 02.24	SE= 00.91
AREA(sq cm)	Mean = 0.001	STD = 0.002	SE= 0.001

Appendix 19c

Pressure/flow graph of 'hamper' for a normal speaker

```
    DUBLIN CHILDREN'S HOSPITAL  SARS REPORT       DATE:01-09-96    By:TRIO
  Measurement Report From File: EK    _HP.R03     Type: HAMPER WITH VOICE
------------------------------------------------------------------------------
Patient:                              Hospital ID:                      Visit:
DX:  0        Age:  0Y   0M         Eval Date: 09/01/96          Examiner: TRIO
------------------------------------------------------------------------------
```



Cur#	VP Area						
1	OP=08.8	NP=-00.1	NFlow=-000.2	SPL=060.0	DP=08.87	AREA= 0.000	
2	OP=07.9	NP=-00.1	NFlow=-000.2	SPL=061.0	DP=07.98	AREA= 0.000	
3	OP=07.7	NP=-00.1	NFlow=001.5	SPL=061.0	DP=07.74	AREA= 0.001	X=4.8
4	OP=08.1	NP=-00.1	NFlow=-000.2	SPL=061.0	DP=08.14	AREA= 0.000	
5	OP=07.4	NP=-00.1	NFlow=003.3	SPL=061.0	DP=07.49	AREA= 0.001	X=6.5
6	OP=08.2	NP=-00.1	NFlow=-000.2	SPL=060.0	DP=08.28	AREA= 0.000	

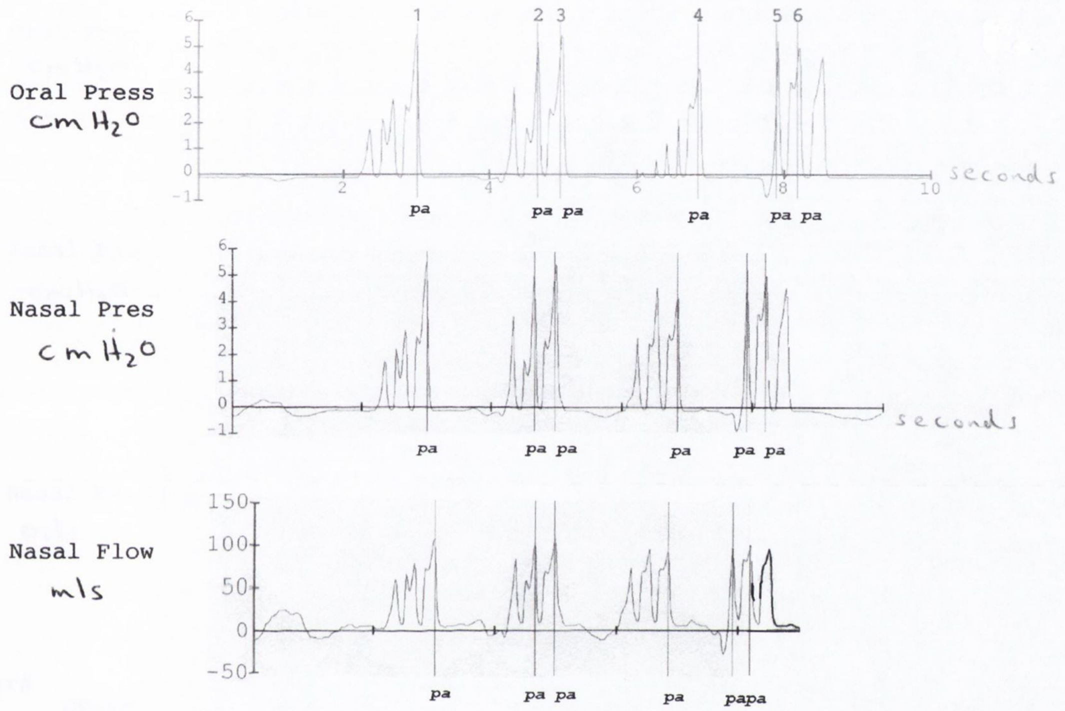
VP Area ===

Oral Press	Mean = 8.0	STD = 00.48	SE= 00.19
Nasal Pres	Mean = -0.1	STD = 00.01	SE= 00.00
Nasal Flow	Mean = 0.7	STD = 01.46	SE= 00.60
SPL	Mean = 60.7	STD = 00.52	SE= 00.21
Computed DP	Mean = 8.08	STD = 00.48	SE= 00.19
AREA(sq cm)	Mean = 0.000	STD = 0.001	SE= 0.000

Appendix 19d

Pressure/flow graph of 'pa' for speaker with Nasal Emission

```
   DUBLIN CHILDREN'S HOSPITAL  SARS REPORT       DATE:05-01-98   By:TRIO
 Measurement Report From File: C      VP.R01     Type: Velopharyngeal Port A
-------------------------------------------------------------------------------
Patient:                             Hospital ID:                      Visit:
DX:  0         Age:  0Y    0M      Eval Date: 11/11/97          Examiner: TRIO
-------------------------------------------------------------------------------
```



Cur#			VP Area			
1	OP=05.5	NP=05.5	NFlow=107.8	DP=00.00	AREA= 0.800	X=3.02
2	OP=05.1	NP=05.1	NFlow=104.4	DP=00.04	AREA= 0.565	X=4.66
3	OP=05.4	NP=05.4	NFlow=107.8	DP=00.02	AREA= 0.800	X=4.97
4	OP=04.1	NP=04.1	NFlow=092.5	DP=00.03	AREA= 0.639	X=6.85
5	OP=05.2	NP=05.2	NFlow=102.7	DP=00.02	AREA= 0.753	X=7.91
6	OP=04.9	NP=04.9	NFlow=101.0	DP=00.06	AREA= 0.452	X=8.20

VP Area

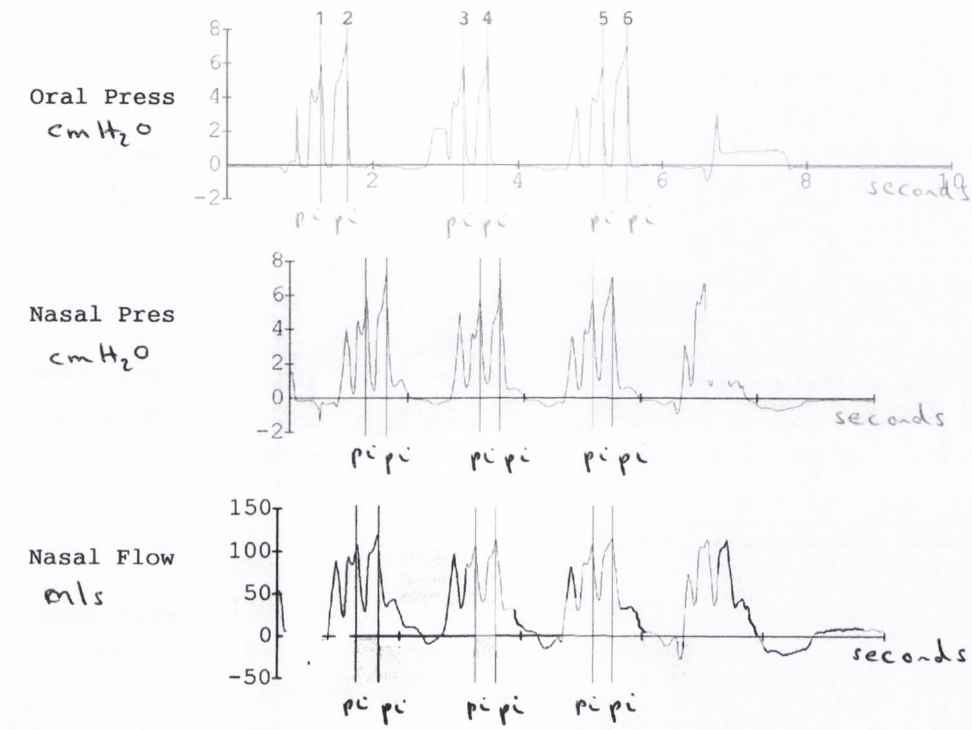
Oral Press	Mean = 5.1	STD = 00.51	SE= 00.21
Nasal Pres	Mean = 5.0	STD = 00.52	SE= 00.21
Nasal Flow	Mean = 102.7	STD = 05.69	SE= 02.32
Computed DP	Mean = 0.03	STD = 00.02	SE= 00.01
AREA(sq cm)	Mean = 0.668	STD = 0.142	SE= 0.058

Appendix 19e

Pressure/flow Graph of 'pi' for speaker with Nasal Emission

DUBLIN CHILDREN'S HOSPITAL SARS REPORT DATE:05-01-98 By:TRIO
Measurement Report From File: C VP.R02 Type: Velopharyngeal Port A

Patient: Hospital ID: Visit:
DX: 0 Age: 0Y 0M Eval Date: 11/11/97 Examiner: TRIO



Cur# ------------------------ VP Area ------------------------

Cur#	OP	NP	NFlow	DP	AREA	X
1	OP=05.2	NP=05.2	NFlow=101.0	DP=00.04	AREA= 0.548	X=1.29
2	OP=07.2	NP=07.2	NFlow=119.7	DP=00.04	AREA= 0.696	X=1.65
3	OP=05.8	NP=05.8	NFlow=107.8	DP=00.02	AREA= 0.800	X=3.24
4	OP=06.9	NP=06.9	NFlow=118.0	DP=00.02	AREA= 0.800	X=3.58
5	OP=05.8	NP=05.8	NFlow=107.8	DP=00.05	AREA= 0.512	X=5.19
6	OP=07.1	NP=07.1	NFlow=118.0	DP=00.02	AREA= 0.800	X=5.52

VP Area =================

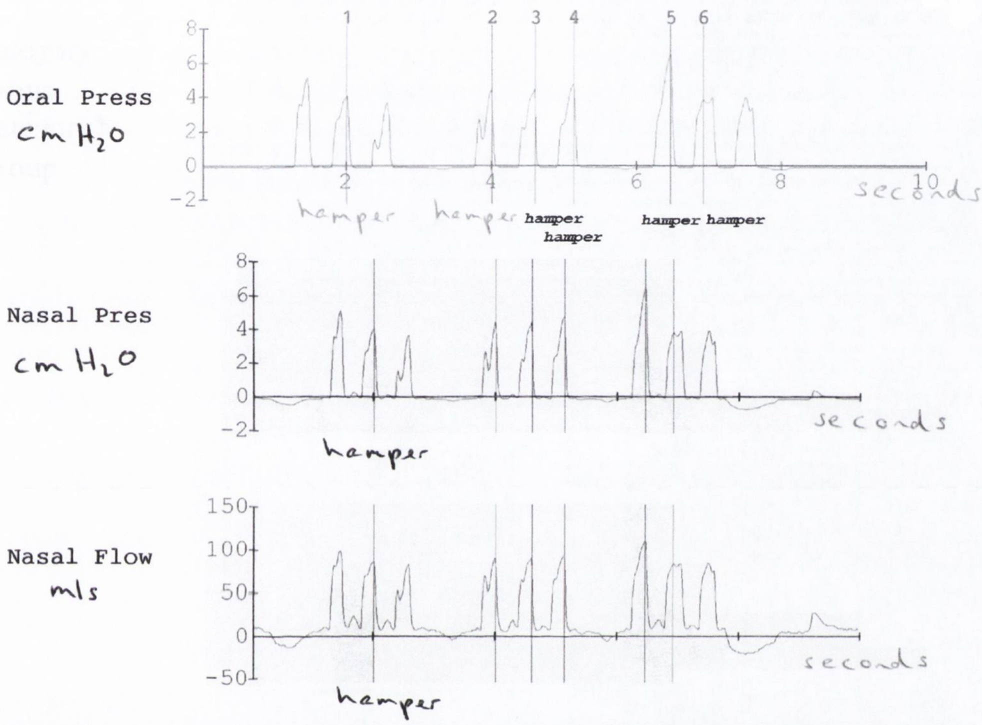
Oral Press	Mean = 6.3	STD = 00.83	SE= 00.34
Nasal Pres	Mean = 6.3	STD = 00.83	SE= 00.34
Nasal Flow	Mean = 112.1	STD = 07.59	SE= 03.10
Computed DP	Mean = 0.03	STD = 00.02	SE= 00.01
AREA(sq cm)	Mean = 0.693	STD = 0.133	SE= 0.054

Appendix 19f

Pressure/flow Graph of 'hamper' for speaker with Nasal Emission

DUBLIN CHILDREN'S HOSPITAL SARS REPORT DATE:05-01-98 By:TRIO
Measurement Report From File: C VP.R03 Type: Velopharyngeal Port A

Patient: Hospital ID: Visit:
DX: 0 Age: 0Y 0M Eval Date: 11/11/97 Examiner: TRIO



Cur#			VP Area			
1	OP=04.0	NP=04.0	NFlow=085.7	DP=00.03	AREA= 0.588	X=2.01
2	OP=04.5	NP=04.4	NFlow=090.8	DP=00.02	AREA= 0.639	X=4.01
3	OP=04.5	NP=04.4	NFlow=090.8	DP=00.04	AREA= 0.481	X=4.61
4	OP=04.8	NP=04.8	NFlow=094.2	DP=00.04	AREA= 0.505	X=5.14
5	OP=06.5	NP=06.4	NFlow=111.2	DP=00.04	AREA= 0.629	X=6.46
6	OP=03.9	NP=03.8	NFlow=084.0	DP=00.03	AREA= 0.573	X=6.92

VP Area

Oral Press	Mean = 4.7	STD = 00.94	SE= 00.38
Nasal Pres	Mean = 4.6	STD = 00.94	SE= 00.38
Nasal Flow	Mean = 92.8	STD = 09.77	SE= 03.99
Computed DP	Mean = 0.03	STD = 00.01	SE= 00.00
AREA(sq cm)	Mean = 0.569	STD = 0.064	SE= 0.026

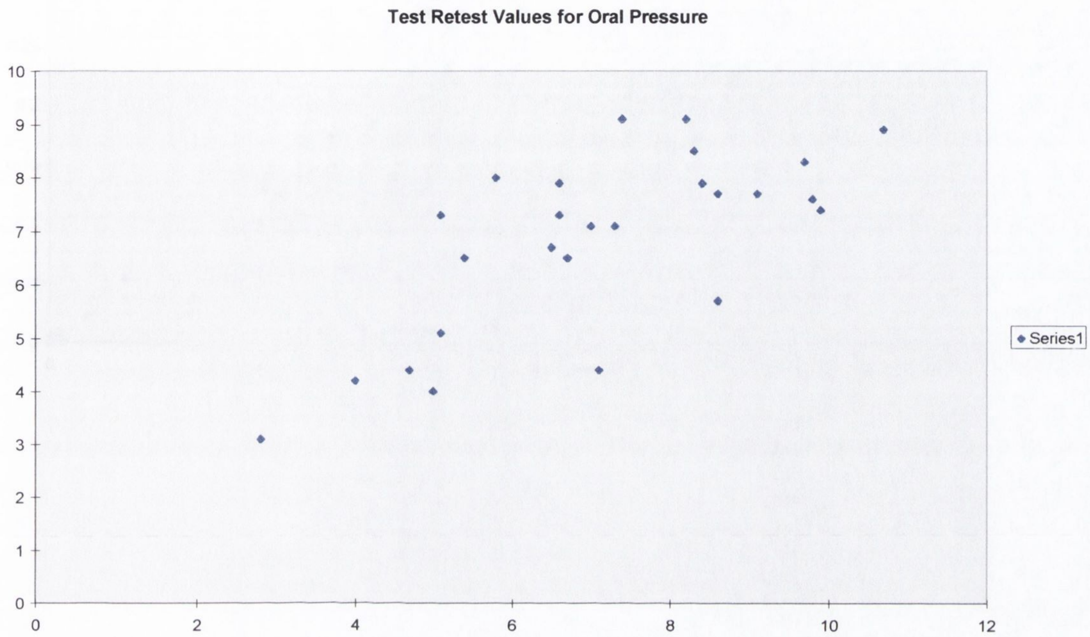
Appendix 20

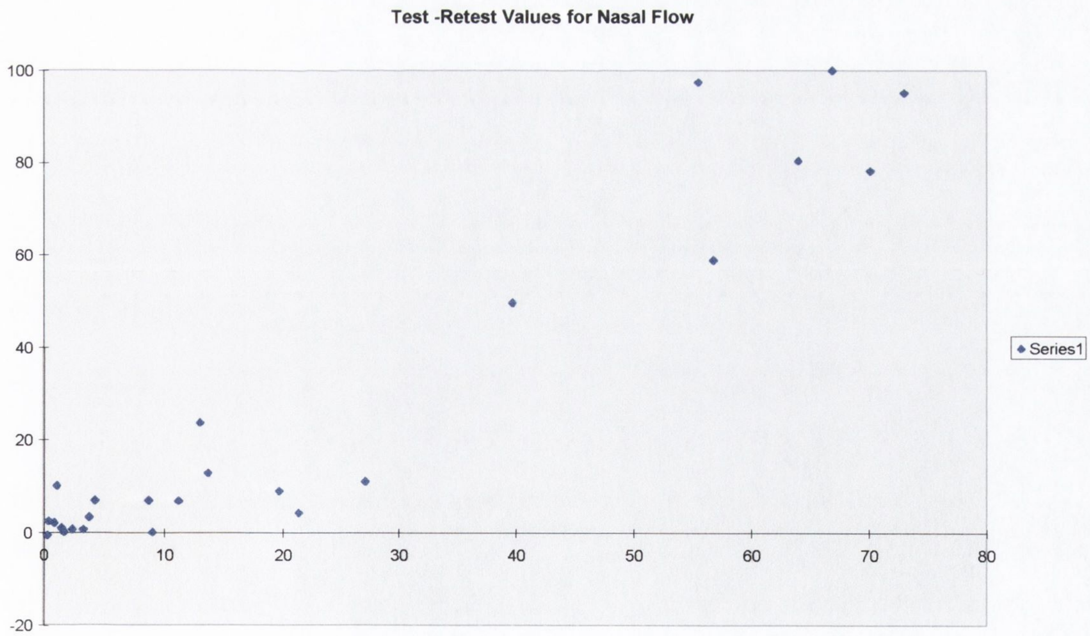
Analysis of Variance For Nasalance Score

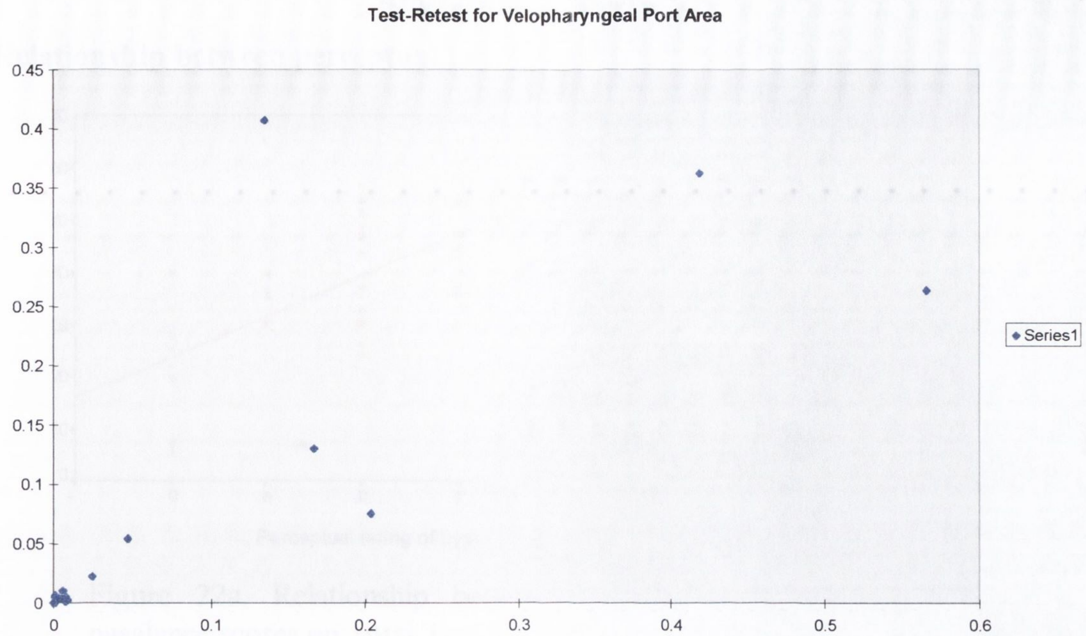
Source	df	Sums of Squares	Mean Square	F-ratio	Prob
Category	4	60033	15008.3	613.6	≤ 0.001
Group	1	24783.1	24783.1	44.1	≤ 0.001
Category* Group	4	6823.81	1705.95	69.7	≤ 0.001

Appendix 21

Pressure/Flow Test-Retest Reliability
X axes indicate measurements for test 1
Y axes indicate measurements for test 2






Test-Retest for Velopharyngeal Port Area
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
0.1
0.2
0.3
0.4
0.5
0.6
Series1

Appendix 22

Relationship between perceptual ratings of hypernasality and nasalance scores.

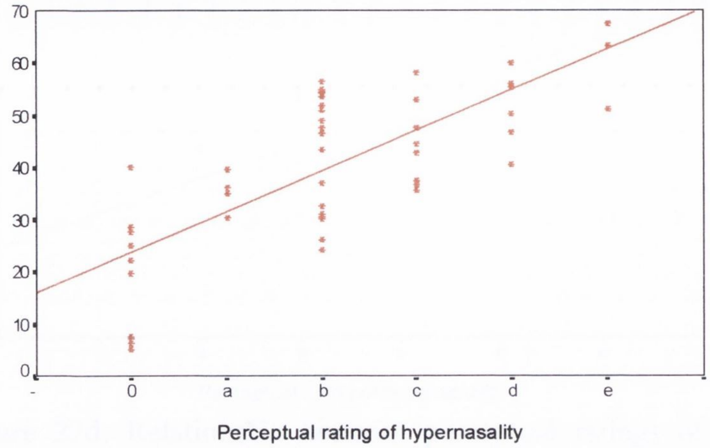


Figure 22a. Relationship between perceptual ratings of hypernasality and nasalance scores on Total Test Sentences. Y axis indicates nasalance scores.

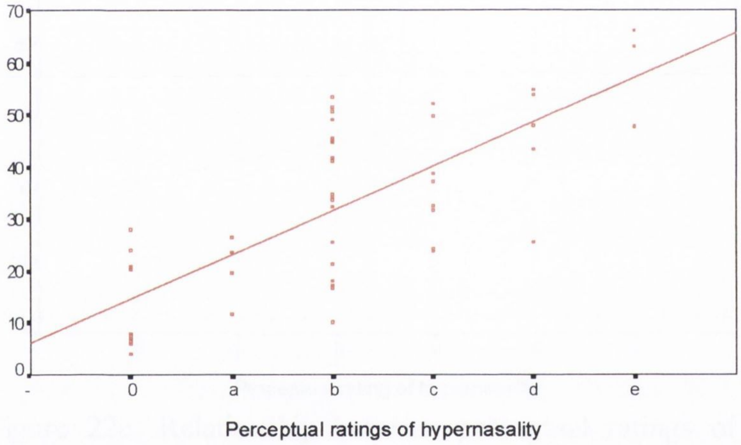


Figure 22b. Relationship between perceptual ratings of hypernasality and nasalance scores on High Pressure Consonant Sentences. Y axis indicates nasalance scores.

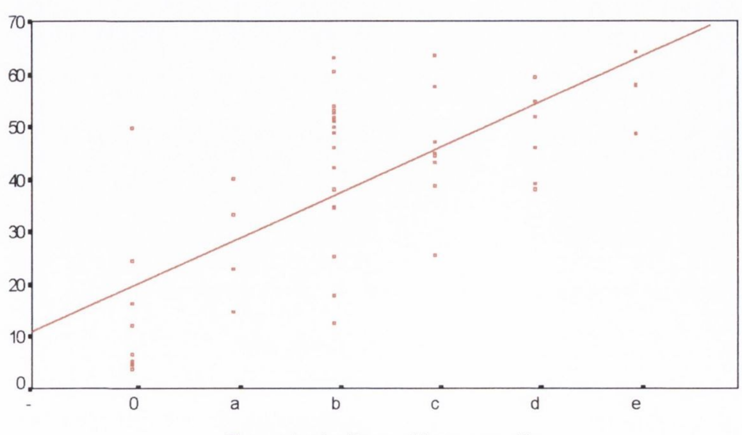


Figure 22c. Relationship between perceptual ratings of hypernasality and nasalance scores on Low Pressure Consonant Sentences. Y axis indicates nasalance scores.

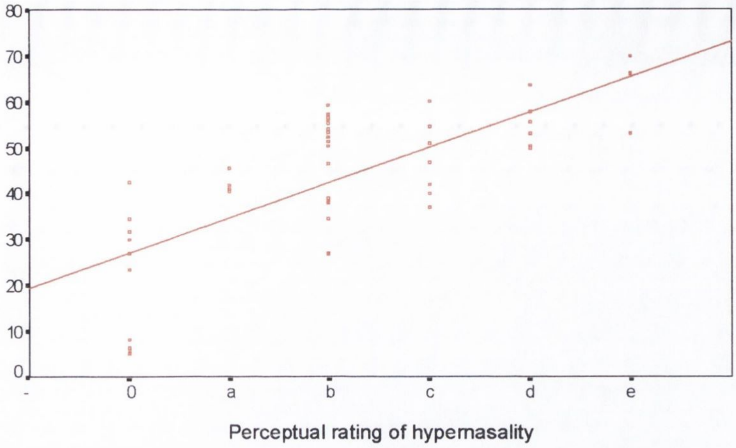


Figure 22d. Relationship between perceptual ratings of hypernasality and nasalance scores on Mixed Consonant Sentences. Y axis indicates nasalance scores.

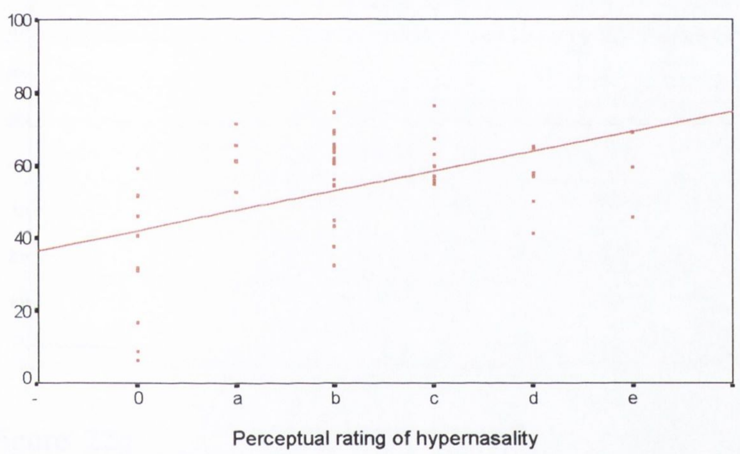


Figure 22e. Relationship between perceptual ratings of hypernasality and nasalance scores on the Nasal Sentence. Y axis indicates nasalance scores.

Relationship between perceptual ratings of hyponasality and nasalance scores

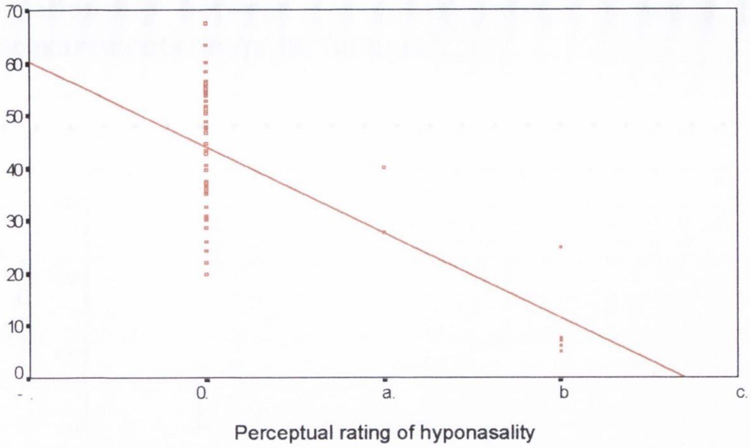


Figure 22f. Relationship between perceptual ratings of hyponasality and nasalance scores on Total Test sentences Y axis indicates nasalance score.

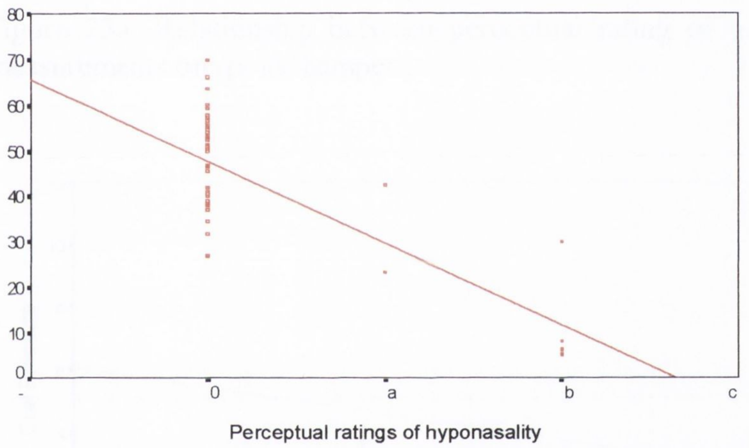


Figure 22g. Relationship between perceptual ratings of hyponasality and nasalance scores on Mixed Consonant sentences. Y axis indicates nasalance score.

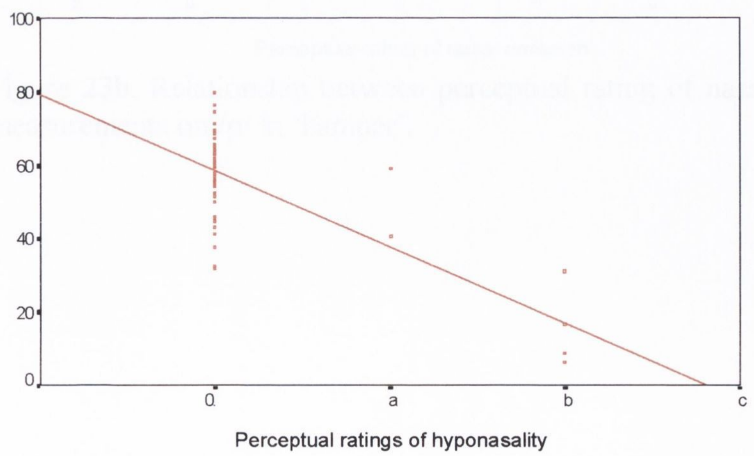


Figure 22h. Relationship between perceptual ratings of hyponasality and nasalance scores on the Nasal sentences. Y axis indicates nasalance score.

Appendix 23

Relationship between perceptual ratings of nasal emission and pressure/flow measurements on /p/ in 'hamper'.

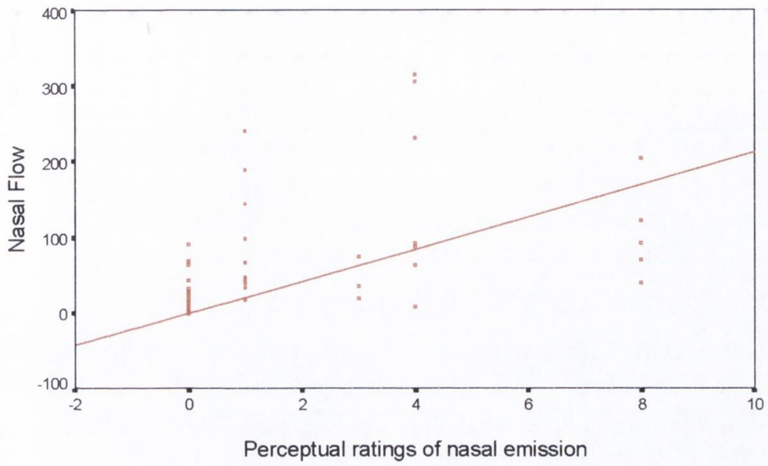


Figure 23a. Relationship between perceptual rating of nasal emission and nasal flow measurements on /p/ in 'hamper'.

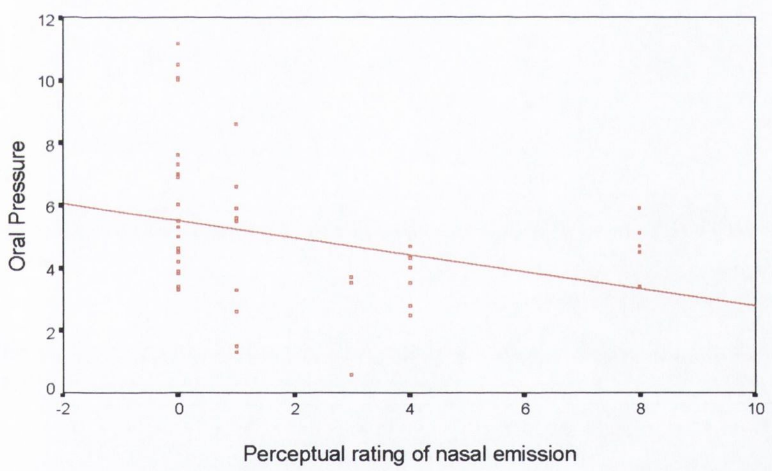


Figure 23b. Relationship between perceptual rating of nasal emission and oral pressure measurements on /p/ in 'hamper'.

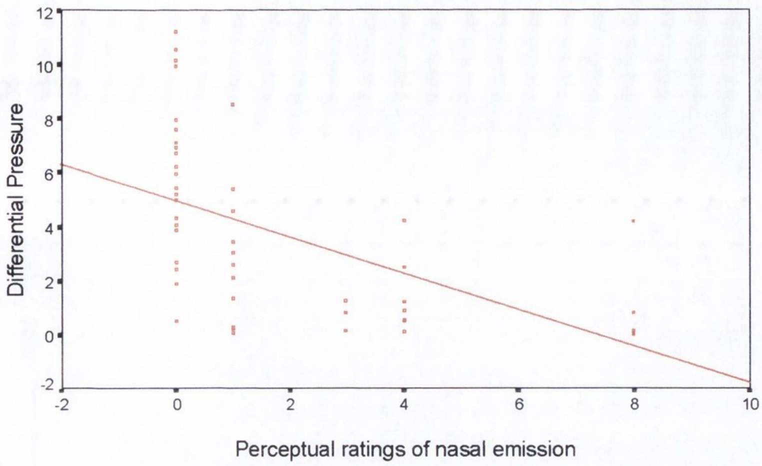


Figure 23c. Relationship between perceptual rating of nasal emission and differential pressure measurements on /p/ in ‘hamper’.

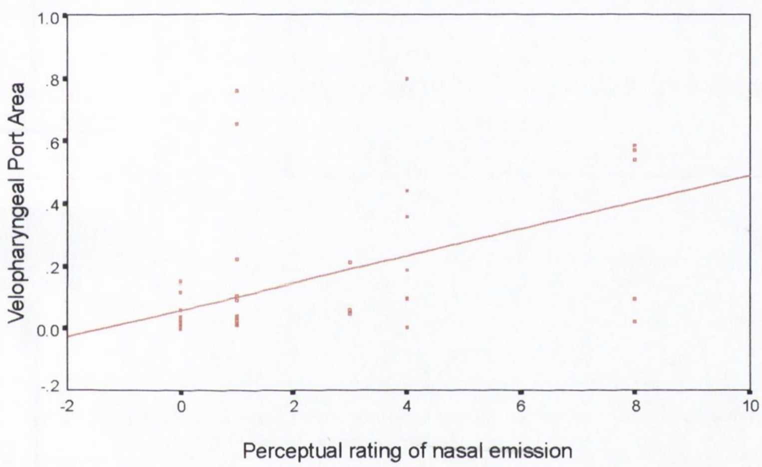
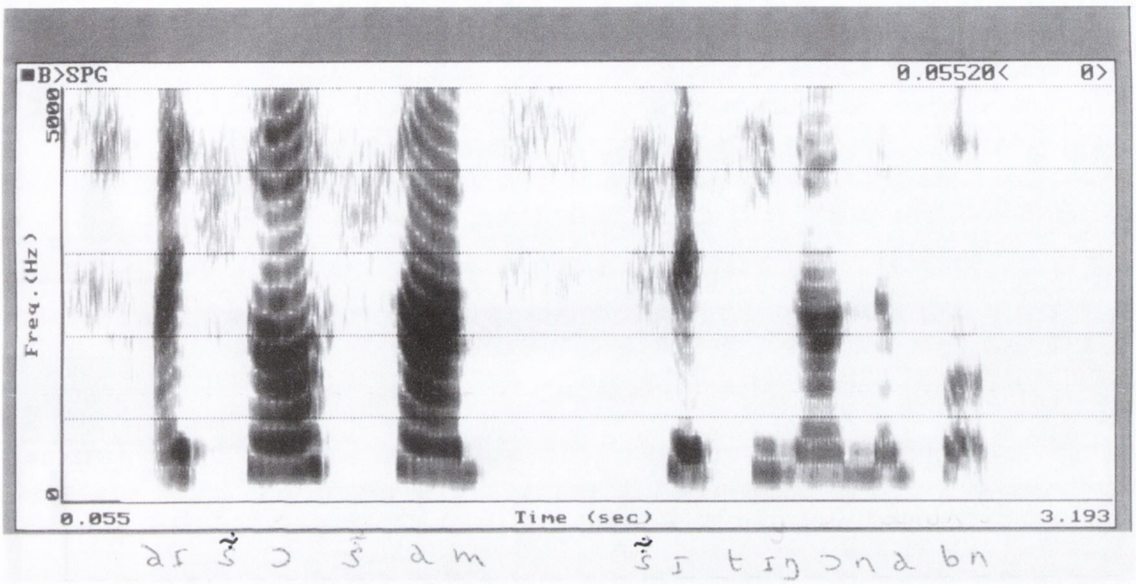
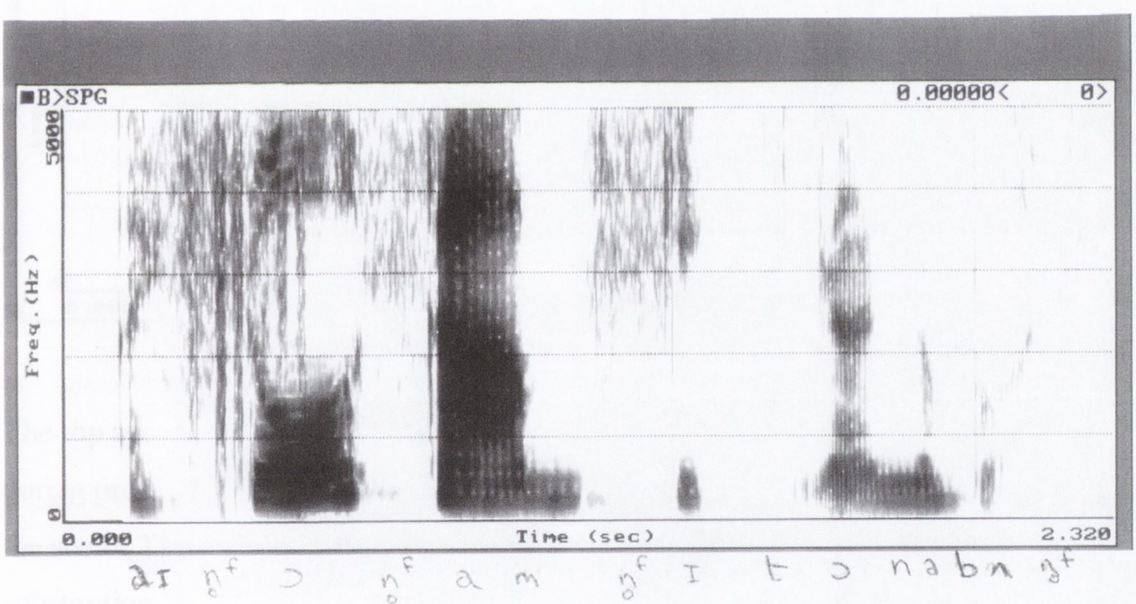


Figure 23d. Relationship between perceptual rating of nasal emission and velopharyngeal port area measurements on /p/ in ‘hamper’.

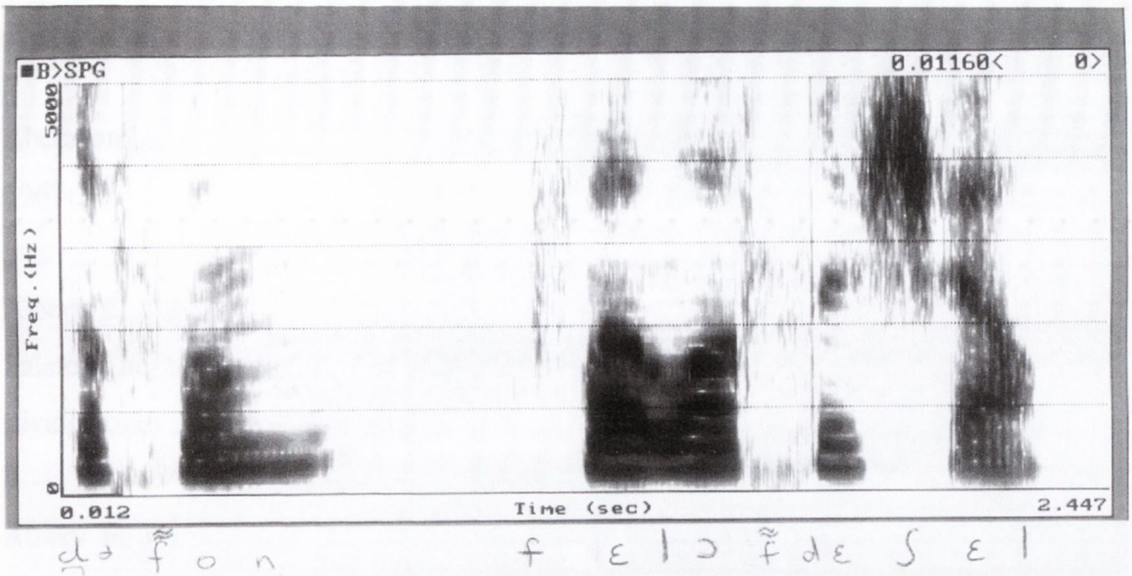
Appendix 24

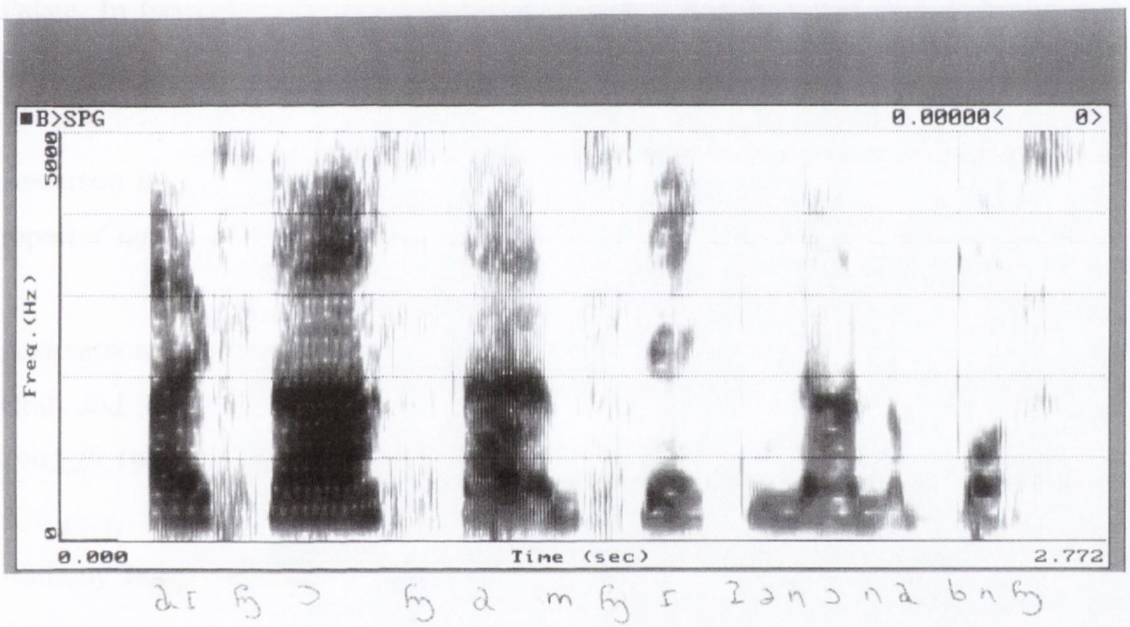
Spectrograms of speakers with nasal emission, nasal fricative, nasal turbulence and velopharyngeal fricative.





The top spectrogram indicates the high frequency energy on the /s̃̈/ sound with nasal emission during the production of the sentence 'I saw Sam sitting on a bus'. Similar high frequency energy is associated with the nasal fricative substitution /$\underset{\circ}{n}^{f}$/ for /s/ in the same sentence.





The top spectrogram indicates the low frequency energy associated with nasal turbulence during production of /f͌/ in the words 'phone' and 'fell' in the sentence 'The phone fell off the shelf'. The bottom spectrogram indicates low frequency energy during the sound substitution /fŋ/ for /s/ in the sentence 'I saw Sam sitting on a bus'.

REFERENCES

Abercrombie D. Elements of General Phonetics. Edinburgh: Edinburgh University Press; 1967

Albery E. Approaches to Treatment of Speech Problems. In Stengelhofen J (ed) Cleft Palate: The Nature and Remediation of Communication Problems. Edinburgh: Churchill Livingstone; 1989.

Albery E, Grunwell P. Consonant Articulation In Different Types Of Cleft Lip And Palate. In Grunwell P. (ed) Analysing Cleft Palate Speech. London: Whurr Publishers Ltd; 1993.

Anderson R. Nasometric Values For Normal Spanish-Speaking Females: A preliminary report. *Cleft Palate Craniofac J* 1996; 33: 333-336.

Andreasson ML, Smith BE, Guyette TW. Pressure-Flow Measurements for Selected Oral and Nasal Sound Segments Produced by Normal Adults. *Cleft Palate J.* 1992;29:19.

Anthony JKF. Aerodynamic and Phonetic Analysis. In Edwards M, Watson ACH (ed) Advances in Management of Cleft Palate. Edinburgh: Churchill Livingstone; 1980.

Bassich CJ, Ludlow CL. The Use of Perceptual Methods By Clinicians For Assessing Voice Quality. *J Speech Hear Dis.* 1986;51:120-125.

Behrman A, Orlikoff R. Instrumentation in Voice Assessment and Treatment: What's the Use? *American Journal of Speech-Language Pathology.* 1997; 4: 9-16.

Borden GJ, Harris KS. Speech Science Primer. Baltimore: Williams & Wilkins; 1980

Bradley DP. Congenital and Acquired Velopharyngeal Inadequacy. In Bzoch KR. (ed) Communication Disorders Related to Cleft Lip and Palate. 3rd edition. Boston: Little Brown & Co; 1989

Bulman JS, Osbourne JF. Statistics in Dentistry. London: British Dental Association, Wimpole Street; 1989.

Bzoch KR. Communication Disorders Related to Cleft Lip and Palate. 3rd edition. Boston: Little Brown & Co; 1989.

Bzoch K, Wood V, Marks R, Williams W, Frovola L, Seagle B, Gainsville FL. Validity Study of a Short Russian Nasometer Test. *Paper presented at American Cleft Palate-Craniofacial Assoc. Conventio., Toronto, Canada.* 1994.

Carney PJ, Sherman D. Severity of Nasality in Three Selected Speech Tasks. *J Speech Hear Res.* 1971;14:396-407.

Clinical Standards Advisory Group. Cleft Lip and/or Palate. London: The Stationary Office; 1998.

Cordez AK. The Reliability of Observational Data: 1 Theories and Methods for Speech-Language Pathology. *J Speech Hear Res.* 1994;37:264-278.

Counihan D. Oral and Nasal Airflow and Air Pressure Measures. In Grabb W, Rosenstein S, Bzoch K. (ed) Cleft Lip and Palate: Surgical, Dental and Speech Aspects. Boston: Little Brown & Co; 1971.

Counihan DT, Cullinan WL. Reliability and Dispersion of Nasality Ratings. *Cleft Palate J.* 1970;7:261-270.

Dalston RM. Using Simultaneous Photodetection and Nasometry to Monitor Velopharyngeal Behaviour During Speech. *J Speech Hear Res.* 1989;32:195-202.

Dalston RM, Neiman GS, Gonzalez-Landa G. Nasometric Sensitivity and Specificity: A Cross-Dialect and Cross-Cultural Study. *Cleft Palate Craniofac J.* 1993; 30: 285-291.

Dalston RM, Seaver EJ. Relative Values of Various Standardised Passages in the Nasometric Assessment of Patients with Velopharyngeal Impairment. *Cleft Palate Craniofac J.* 1992; 29:17-21.

Dalston R., Warren D. Comparison of Tonar ll, Pressure-Flow, and Listener Judgements of Hypernasality in Assessment of Velopharyngeal Function. *Cleft Palate J.* 1986; 23:108-115

Dalston RM, Warren D, Dalston E. A Preliminary Investigation Concerning Use of Nasometry in Identifying Patients with Hyponasality and/or Nasal Airway Impairment. *J Speech Hear Res.* 1991a; 34:11-18.

Dalston RM, Warren D, Dalston E. Use of Nasometry as a Diagnostic Tool for Identifying Patients with Velopharyngeal Impairment. *Cleft Palate Craniofac. J* 1991b;28;1184-188.

Dalston RM, Warren D, Dalston E. The Temporal Characteristics of Aerodynamic Phenomena Associated with Patients Manifesting Varying Degrees of Velopharyngeal Adequacy. *Folia Phoniat.* 1991c;43:226-233.

Dalston RM, Warren DW, Morr KE, Smith LR. Intraoral Pressure and Its Relationship to Velopharyngeal Inadequacy. *Cleft Palate J.* 1988;25:210-219.

D'Antonio LL, Muntz H, Province M. Laryngeal/Voice Findings in Patients with Velopharyngeal Dysfunction. *Laryngoscope.* 1988;98:432-438.

D'Antonio LL, Scherer NJ. The Evaluation of Speech Disorders Associated with Clefting. In Shprintzen RJ, Bardach J. (ed) Cleft Palate Speech Management: A Multidisciplinary Approach. St Louis, Missouri: Mosby-Year Book, Inc; 1995.

Duckworth M, Allen G, Hardcastle W, Ball M. Extension to the International Phonetic Alphabet for the Transcription of Atypical Speech. *Clinical Linguistics and Phonetics.* 1990;4:273-280.

Ellis RE, Flack FC, Curle HJ, Selly WG. A System for Assessment of Nasal Airflow During Speech. *British Journal of Disorders of Communication.* 1978; 13:31-40.

Eurocleft Speech Group: Brondsted K, Grunwell P, Henningsson G, Jansonius K, Karling J, Meijer M, Ordin U, Wyatt R, Vermeij-Zieverink E.
Cleft Speech in a European Perspective: Eurocleft Speech Project. In Grunwell P. (ed) Analysing Cleft Palate Speech. London: Whurr Publishers Ltd; 1993.

Fairbanks G. Voice and Articulation Drillbook. New York: Harper Row; 1960.

Fleiss JL. Statistical Methods for Raters and Proportions. New York: John Wiley & Sons Inc; 1981.

Fletcher SG. Contingencies for Bioelectronic Modification of Nasality. *J Speech Hear Dis.* 1972; 37:329-346.

Fletcher SG. Nasalance vs. Listener Judgements of Nasality. *Cleft Palate J.* 1976; 13:31-44 .

Fletcher SG, Adams LE, McCrutcheon MJ. Cleft Palate Speech Assessment Through Oral-Nasal Acoustic Measures. In Bzoch KR. (ed) Communicative Disorders related to Cleft Lip and Palate. Boston: Little-Brown; 1989.

Fletcher SG, Bishop M. Measurement of Nasality with Tonar. *Cleft Palate J.* 1970; 7: 610 - 621.

Folkins JW, Moon JB. Approaches to the Study of Speech Production. In Bardach J, Morris HL. (ed) Multidisciplinary Management of Cleft Lip and Palate. Philadelphia: W.B. Saunders & Co; 1990.

Fujiwara Y, Hiramoto M, Kawano M. Assessment of Velopharyngeal Function: Pneumotography and Videonasofiberscopy. *Paper presented at 7th International Congress on Cleft Palate and Craniofacial Anomalies. Australia*. 1993.

Geddes LA, Baker LE. Principles of Applied Biomedical Instrumentation. 3rd edition. London: John Wiley & Sons; 1989.

Gerratt B, Kreiman J, Antonanzas-Barroso N, Berke GS. Comparing Internal and External Standards in Voice Quality *J Speech Hear Res.* 1993; 36:14-20.

Grunwell P, Harding A. PACSTOYS *A Screening Assessment of Phonological Development* Windsor NFER-Nelson. 1995.

Grunwell P, Harding A. Describing Types of Nasality. *Clinical Linguistics and Phonetics.* 1996;10: 157-161.

Grunwell P, Sell D, Harding A. Describing Cleft Palate Speech. In Grunwell P. (ed) Analysing Cleft Palate Speech. London: Whurr Publishers Ltd; 1993.

Haapanen ML. Nasalance Scores in Normal Finnish Speech. *Folia Phoniatrica* 1991a; 43: 197-203.

Haapanen ML. A Simple Clinical Method of Evaluating Perceived Hypernasality. *Folia Phoniatrica.* 1991b; 43: 122-132.

Hamlet S. Vocal Compensation: An Ultrasonic Study of Vocal Fold Vibration in Normal and Nasal Vowels. *Cleft Palate J.* 1973;26:267-259.

Hardin MA, Van Demark DR, Morris HL, Payne MM. Correspondence between Nasalance Scores and Listener Judgements of Hypernasality and Hyponasality. *Cleft Palate-Craniofac J.* 1992;29: 346-351.

Harding A, Harland K, Razzell R. Cleft Audit Protocol for Speech (CAPS). *Speech/Language Therapy Dept., Broomfield Hospital, Court Road, Chelmsford, Essex, UK.* 1997.

Harding A, Grunwell P Active Versus Passive Cleft-Type Speech Characteristics. *International Journal of Language & Communication Disorders.* 1998; 33:329-352.

Harkins CS, Berlin A, Harding R, Longacre JJ, Snodgrasse R. Report of the Nomenclature Committee of the American Cleft Palate Association. *Cleft Palate Bulletin* 1960; 10:11.

Henningsson G, Hutters B. Perceptual Assessment of Cleft Palate Speech- With Special Reference to Minimum Standards For Inter-Centre Comparisons Of Speech Outcomes. *Unpublished Proposal for 8th International Congress on Cleft Palate and Related Craniofacial Anomalies. Singapore*. 1997.

Hutters B, Brondsted K. Strategies in Cleft Palate Speech - with Special Reference to Danish. *Cleft Palate J.* 1987; 24:126-136.

Hutters B, Brondsted K. A Simple Nasal Anemometer for Clinical Purposes. *European J of Disorders of Communication.* 1992; 27: 101-119.

Hyman L. Phonology Theory and Analysis. New York: Holt, Rinehart and Winston; 1975.

Isshiki N, Honjow I, Morimoto M. Effects of Velopharyngeal Incompetence upon Speech. *Cleft Palate J.* 1968; 5:297-310.

Karnell MP. Discrimination of Hypernasality and Turbulent Nasal Airflow. *Cleft Palate-Craniofac J* 1995; 32: 145-148.

Karnell MP, Folkins JW, Morris HL. Relationships between the Perception of Nasalization and Speech Movements in Speakers with Cleft Palate. *J Speech Hear Res.* 1985;28:63-72.

Kearns KP, Simmons NN. Interobserver Reliability and Perceptual Ratings: More Than Meets The Ear. *J Speech Hear Res.* 1988;31:131-136.

Kent RD. Hearing and Believing: Some Limits to the Auditory-Perceptual Assessment of Speech and Voice Disorders. *American Journal of Speech and Language Pathology.* 1996; 5: 7-23.

Kittelson C, Broen P, Moller K. Procedures and Symbols for the Transcription of the Speech of Children and Adults with Cleft Palate. (Unpublished paper) University of Minnesota. 1983.

Kreiman J, Gerratt B, Kempster GB, Berke GS. Perceptual Evaluation of Voice Quality: Review, Tutorial, and a Framework for Future Research. *J Speech Hear Res* 1993; 36:21-40.

Kreiman J, Gerratt B, Precoda K. Listener Experience and Perception of Voice Quality. *J Speech Hear Res* 1990; 33:103-115.

Kuehn DP. Assessment of Resonance Disorders. In Lass NJ, McReynolds LV, Northam JC, Yoder DE. (ed) Speech, Language and Hearing. Vol. 11, Pathologies of Speech and Language. Phil.:WB Saunders Co; 1982.

Kummer AW, Curtis C, Wiggs M, Lee L, Strife JL. Comparison of Velopharyngeal Gap Size in Patients with Hypernasality, Hypernasality and Nasal Emission, or Nasal Turbulence (Rustle) as the Primary Speech Characteristic. *Cleft palate - Craniofac J.* 1992;29:152-156.

LaBlanc GR, Steckol KF, Cooper MH. Advances in Non Invasive Measures of Vocal Acoustics. *Ear, Nose and Throat Journal*. 1991; 70: 678-684.

Ladefoged P. Instrumental Techniques for Linguistic Phonetic Fieldwork. In Hardcastle WJ, Laver J (ed) The Handbook of Phonetic Sciences. Oxford: Blackwell Publishers; 1997.

Laine T, Warren DW, Dalston RM, Morr KE. Screening of Velopharyngeal Closure Based on Nasal Airflow Rate Measurements. *Cleft Palate J.* 1988;25: 220-225.

Laver J. The Phonetic Description of Voice Quality. Cambridge: Cambridge University Press. 1980.

Laver J, Wirz S, Mackenzie Beck J, Hiller S. A Perceptual Protocol for the Analysis of Vocal Profiles. Edinburgh: University of Edinburgh (Dept. of Linguistics). Work in Progress. 1981.

LeBlanc EM, Shprintzen RJ. The Velopharyngeal Mechanism. In Berkowitz S.(ed) Cleft Lip and Palate: Perspectives in Management. Vol 11. San Diego:Singular Publishing Group Inc; 1996.

Leder S, Lerman J. Some Acoustic Evidence for Vocal Abuse in Adult Speakers with Repaired Cleft Palate. *Laryngoscope.* 1985;95:837-840.

Lewis J, Andreassen M, Leeper H, Macrae D. Vocal Characteristics in Children with Cleft Lip/Palate and Associated Velopharyngeal Incompetence. *J of Otolaryhgology.* 1993;22:113-117.

Lintz LB, Sherman D. Phonetics Elements and Perception of Nasality. *J Speech Hear Res.* 1961;4:381-396.

Litzaw LL, Dalston RM. The Effect of Gender upon Nasalance Scores among Normal Adult Speakers. *J Commun Disorders* 1992;25:55-64.

Lohmander-Agerskov A. Speech in Children with Cleft Lip and Palate treated with Delayed Closure of the Hard Palate. PhD Thesis. Department of Logopedics and Phoniatrics, Sahlgrenska University Hospital, Goteborg University, Sweden. 1996.

Loney RW, Bloem TJ. Velopharyngeal Dysfunction: Recommendations for Use of Nomenclature. *Cleft Palate Journal.* 1987; 24: 334-335.

Main A, Kelly S, Manley G. Instrumental Assessment and Treatment of Hypernasality, following Maxillofacial Surgery, using SNORS: a single case study. *International Journal of Language & Communication Disorders*. 1999; 34: 223-238

Mackay IR, Kummer AW. Simplified Nasometric Assessment Procedures. Lincoln Park, New Jersey: Kay Elemetrics; 1994.

Mayo R, Floyd LA, Warren D, Dalston R, Mayo CM. Nasalance and Nasal Area Values: Cross-Racial Study. *Cleft Palate-Craniofac J.* 1996;33: 143- 149.

McWilliams BJ, Glaser ER, Phillips BJ, Lawrence C, Lavorato AS, Berry QC, Skolnick ML. A Comparative Study of Four Methods of Evaluating Velopharyngeal Adequacy. *Plastic Reconstr. Surg.* 1981;68: 1-10.

McWilliams BJ, Philips BJ. Velopharyngeal Incompetence. An Audio Seminar. Philadelphia: B.C. Decker Inc; 1979.

McWilliams BJ, Morris H, Shelton R. Cleft Palate Speech. 2nd edn. Philadelphia: BC Decker Inc; 1990.

Mirlohi HR, Kelly SW, Manley MCG. New Technique for Assessment of Velopharyngeal Function. *Medical and Biological Engineering and Computing* 1994;32:562-566.

Moll KL. Objective Measures of Nasality. *Cleft Palate J.* 1964;1:371-374.

Moller KT, Starr CD. The Effects of Listening Conditions on Speech Ratings Obtained in a Clinical Setting. *Cleft Palate J.* 1984;21:65-69.

Moon JB. Evaluation of Velopharyngeal Function. In Moller KT & Starr CD. (ed) Cleft Palate: Interdisciplinary Issues and Treatment for Clinicians by Clinicians. Texas: Pro-Ed Inc; 1993.

Morr K, Warren DW, Dalston RM, Smith LR. Screening of Velopharyngeal Inadequacy by Differential Pressure Measurements. *Cleft Palate J.* 1989;26: 42-45.

Morris HL. Types of Velopharyngeal Incompetence. In Winitz H. (ed) Treating Articulation Disorders; for Clinicians by Clinicians. Baltimore: University Park Press; 1984.

Morris HL, Spriesterbach DC, Shelton RL. An Articulation Test for Assessing Competency of Velopharyngeal Closure. *J Speech Hear Res.* 1961;4:48-55.

Nasometer Model 6200-3. Instruction Manual. New Jersey: Kay Elemetrics Corporation; 1994

Nellis JL, Neiman GS, Lehman JA. Comparison of Nasometry and Listener Judgements of Nasality in Assessment of Velopharyngeal Function After Pharyngeal Flap Surgery. *Cleft Palaet- Craniofac J* 1992; 29 : 157-163.

Nichols AC. Nasalance Statistics for Two Mexican Populations. *Cleft Palate-Craniofac J* 1999;36:57-63.

O'Neill L, Malone JF. Physiological Measures with Transducers. In McMullen JT (ed) Physical Techniques in Medicine. Vol. 11. London: John Wiley & Sons; 1977.

Painter C. An Introduction to Instrumental Phonetics. Baltimore: University Park Press; 1979.

Parker AJ, Maw AR, Szallasi F. An Objective Method of Assessing Nasality: A Possible Aid in the Selection of Patients for Adenoidectomy. *Clin. Otolaryngol.* 1989;14: 161-166

Paynter ET, Watterson TL, Boose WT. The Relationship between Nasalance and Listener Judgements. *Paper presented at American Cleft Palate-Craniofacial Assoc. Conventio., Hilton Head, South Carolina.* 1991.

PERCI Speech-Aeromechanics Research System. System Manual. Chapel Hill, N.C: Microtronics Corporation; 1994

Peterson S. Nasal Emission as a Component of the Misarticulation of Sibilants and Affricates. *J Speech Hear D.*1975; 40:106-114.

Philips BJ. Speech Assessment. *Semin Speech Language.* 1986; 55:297.

Philips BJ, Bzoch KR. Reliability of Judgements of Articulation of Cleft Palate Speakers. *Cleft Palate J.*1969;6:24-34.

Pinborough-Zimmerman J, Canady C, Yamashiro DK, Morales L. Articulation and Nasality Changes Resulting from Sustained Fistula Obturation. *Cleft palate-Craniofac J.* 1998;35:81-87.

Raaijmakers MF, Dekker J, Dejonckere PH, van der Zee J. Reliability of the Assessment of Impairment, Disabilities and Handicaps in Survey Research on Speech Therapy. *Folia Phoniatr.* 1995;47:199-209.

Ramig LA. Effects of Examiner Expectancy on Speech Ratings of Individuals with Cleft Lip and/or Palate. *Cleft Palate J.* 1982;19:270-274.

Rampp D, Counihan D. Vocal Pitch-Intensity Relationships in Cleft Palate Speakers. *Cleft Palate J.* 1970;7:846-857.

Riski JE, Warren DW, Zajac DJ, Lutz RE. PERCI-SARS Aerodynamic Assessment of Speech. *Study Session at American Cleft Palate Craniofacial Assoc. Tampa, Florida.* 1995.

Sapienza C, Brown W, Williams W, Wharton P, Turner G. Respiratory and Laryngeal Function Associated with Experimental Coupling of the Oral and Nasal Cavities. *Cleft Palate-Craniofac J.* 1996;33:118-125.

Schiavetti N. Scaling Procedures for the Measurement of Speech Intelligibility. In Kent RD. (ed) Intelligibility in Speech Disorders. Amsterdam: John Benjamins Publ. Co; 1992.

Seaver EJ, Dalston RM, Leeper HA, Adams LE. A Study of Nasometric Values for Normal Nasal Resonance. *J Speech Hear Res* 1991; 34: 715-721.

Sell D. An Exploration of the UK CSAG Speech Study - What have we learnt? *Paper presented at 6th European Craniofacial Meeting, Manchester. June 1999.*

Sell D, Grunwell P. Speech Results Following Late Palatal Surgery in Previously Unoperated Sri Lankan Adolescents With Cleft Palate. *Cleft Palate J.* 1990; 27:162-168.

Sell D, Grunwell P. Speech in Subjects with Late Operated Cleft Palate. In Grunwell P. (ed) Analysing Cleft Palate Speech. London: Whurr Publishers Ltd; 1993.

Sell D, Grunwell P. Speech Assessment and Therapy, In Watson T, Grunwell P, Sell D. (ed) Management of Cleft Lip and Palate. London: Whurr Publishers Ltd; 2000.

Sell D, Harding A, Grunwell P. A Screening Assessment of Cleft Palate Speech (Great Ormond Street Speech Assessment). *European J Disorders of Communication* 1994;29:1-15.

Sell D, Harding A, Grunwell P. GOS.SP.ASS. '94 and '98 - A Training Video of Speech Characteristics. Institute of Child Health and Great Ormond Street Hospital for Children NHS Trust. 1998.

Sell D, Harding A, Grunwell P. GOS.SP.ASS'98 An Assessment for Speech Disorders Associated with Cleft Palate and/or Velopharyngeal Dysfunction (revised). *International Journal of Language & Communication Disorders* 1999;34: 17-33.

Sell D, Sweeney T. Response to Instrumental Assessment and Treatment of Hypernasality, following Maxillofacial Surgery, using SNORS: A Single Case Study. *IJDC.* 1999:34: 223-238. Main A, Kelly S, Manley G. *IJDC* (in press).

Sherman D. The Merits of Backward Playing Of Connected Speech In the Scaling Of Voice Quality Disorders. *J Speech Hear D* 1954;19:312-321.

Shprintzen R. Instrumental Assessment of Velopharyngeal Valving. In Shprintzen RJ, Bardach J. (ed) Cleft Palate Speech Management: A Multidisciplinary Approach. St. Louis: Mosby-Year Book Inc; 1995.

Smith BE, Guyette TW. Pressure-Flow Differences in Performance During Production of the CV Syllables /pi/ and /pa/. *Cleft Palate-Craniofac J.* 1996;33:74-76.

Smith BE, Skef Z, Cohen M, Dorf DS. Aerodynamic Assessment of the Results of Pharyngeal Flap Surgery: A Preliminary Investigation. *Plastic Reconstr Surg.* 1985;76:402-408.

Spriesterbach DC, Power GR. Nasality in Isolated Vowels and Connected Speech of Children with Cleft Palate. *J Speech Hear Res.* 1959;2:40-45.

Stengelhofen J. Working with Cleft Palate. Bicester. UK: Winslow Press; 1990.

Streiner DL & Norman GR. Health Measurement Scales: A practical guide to their development and use. (2nd edt) UK: Oxford University Press; 1995.

Swartz M. Acoustic measures of Nasalization and Nasality. In Grabb W, Rosenstien S, Bzoch K (ed) Cleft Lip and Palate: Surgical, Dental & Speech Aspects. Boston: Little Brown & Co; 1971.

Sweeney T. Assessment of Nasality in Pathological Speech. MSc Thesis, University of Edinburgh. 1984.

Sweeney T, Sell D, Grunwell P. Objective Measures of Nasality and Nasal Airflow. *Paper presented at Annual Conference of Craniofacial Society of Great Britain. Egham.* 1996.

Sweeney T, Sell D, Grunwell P. The Relationship between Perceptual and Instrumental Measurements of Nasality and Nasal Airflow. *Paper presented at the 6th European Craniofacial Meeting, Manchester.* June 1999.

Tarlow A, Saxman J. A Comparative Study of the Speaking Fundamental Frequency Characteristics in Children with Cleft Palate. *Cleft Palate J.* 1970;7:696-705.

Thompson AE, Hixon TJ. Nasal Air Flow During Normal Speech Production. *Cleft Palate J.* 1979;16:412-420.

Trindade IEK, Genero KF, Dalston RM. Nasalance Scores of Normal Brazilian Portuguese Speakers. *Brazilian J of Dysmorphology & Speech-Hearing Disorders* 1997; 1:23-34.

Trost JE. Articulatory Additions to the Classic Description of the Speech of Persons with Cleft Palate. *Cleft Palate J.* 1981;18:193-198.

Vallino-Napoli L, Montgomery AA. Examination of the Standard Deviation of Mean Nasalance Scores in Subjects with Cleft Palate: Implications for Clinical Use. *Cleft Palate-Craniofac J.* 1997; 34: 512-519.

Van Demark D. Misarticulation and Listener Judgements of the Speech of Individuals with Cleft Palates. *Cleft Palate J.* 1964;1:232-245.

Van Demark D, Bzoch K, Daly D, Fletcher S, McWilliams BJ, Pannbacker M. and Weinberg B. Methods of Assessing Speech in Relation to Velopharyngeal Function. *Cleft Palate J.* 1986; 23: 215-224.

van Doorn J, Purcell A. Nasalance Levels in the Speech of Normal Australian Children. *Cleft Palate-Craniofac J.* 1998; 35:287-292.

Van Hattum RJ, Worth JH. Airflow Rates in Normal Speakers. *Cleft Palate J.* 1967;4:137-147.

Warren DW. Nasal Emission of Air and Velopharyngeal Function. *Cleft Palate J.* 1967;4: 148-156.

Warren DW. The Determination of Velopharyngeal Incompetence by Aerodynamic and Acoustical Techniques. *Clinics in Plastic Surgery.* 1975;2:299-304.

Warren DW. Perci: A Method for Rating Palatal Efficiency. *Cleft Palate-Craniofac J* 1979;16:279-285.

Warren DW. Compensatory Speech Behaviour in Individuals with Cleft Palate: a regulation/control phenomenon? *Cleft Palate J* 1986;23:251-260.

Warren DW. Aerodynamic Assessment of Velopharyngeal Performance. In Bzoch KR. Communication Disorders Related to Cleft Lip and Palate. 3rd edition. Boston: Little Brown & Co; 1989.

Warren DW, Dalston RM, Mayo R. Aerodynamics of Nasalization. *Phonetics and Phonology.* 1993a; 5: 119-145.

Warren DW, Dalston RM, Mayo R. Hypernasality in the Presence of 'Adequate' Velopharyngeal Closure. *Cleft Palate-Craniofac J.* 1993b;30:150-154.

Warren DW, Dalston RM, Mayo R. Hypernasality and Velopharyngeal Impairment. *Cleft Palate-Craniofac J.* 1994;31:257-262.

Warren DW, Dalston RM, Morr KE, Hairfield WM, Smith LR. The Speech Regulating System: Temporal and Aerodynamic Responses to Velopharyngeal Inadequacy. *J Speech Hear Res.* 1989;32:566-575.

Warren DW, Dalston RM, Trier WC, Holder MB. A Pressure-Flow Technique for Quantifying Temporal Patterns of Palatopharyngeal Closure. *Cleft Palate J.* 1985;22:11-19.

Warren DW, Duany LF, Fuscher ND. Nasal Pathway Resistance in Normal and Cleft Lip and Palate Subjects. *Cleft Palate J.* 1969;6:134-140.

Warren DW, Dubois AB. A Pressure - Flow Technique for Measuring Velopharyngeal Orifice Area During Continuous Speech. *Cleft Palate J.* 1964;1:52-71.

Watson ACH. Classification of Cleft Lip and Palate. In Watson T, Grunwell P, Sell D. (ed) Management of Cleft Lip and Palate. London: Whurr Publishers Ltd; 2000.

Watterson T, Hinton J, McFarlane S. Novel Stimuli for Obtaining Nasalance Measures In Young Children. *Cleft Palate-Craniofac J.* 1996;33:67-73.

Watterson T, Lewis KE., Dalston RM. Interpretation and Clinical Application of Nasalance Scores. *Paper presented at the Annual Conference of the American Speech, Language and Hearing Association.* Nov. 1998b.

Watterson T, Lewis KE, Deutsch C. Nasalance and Nasality in Low Pressure and High Pressure Speech. *Cleft Palate-Craniofac J.* 1998a; 35; 293-298.

Watterson T, Lewis KE, Foley-Homan N. Effect of Stimulus Length on Nasalance Scores. *Cleft Palate-Craniofac J.* 1999;36:243-247.

Watterson T, McFarlane S, Wright DS. The Relationship Between Nasalance and Nasality in Children with Cleft Palate. *Journal of Communication Disorders.* 1993; 26:13-28.

Williams RG, Eccles R, Hutchings H. The Relationship between Nasalance and Nasal Resistance to Airflow. *Acta Otolaryngol Stock* 1990;110:443-449.

Wilson DK. Voice Problems of Children. 2nd edition. Baltimore: Williams & Wilkins; 1987.

Wirz S, Mackenzie Beck J. Assessment of Voice Quality: the Vocal profiles Analysis Scheme. In: Wirz S.(ed) Perceptual Approaches to Communication Disorders. London: Whurr Pub Ltd; 1995.

Wise CM. Applied Phonetics. Englewood Cliffs, New Jersey: Prentice Hall; 1957.

Witzel MA. Communication Impairment Associated with Clefting. In Shprintzen RJ, Bardach J. (ed) Cleft Palate Speech Management: A Multidisciplinary Approach. St. Louis: Mosby-Year Book Inc; 1995.

Wyatt R, Sell D, Russell J, Harding A, Albery E. Cleft Palate Speech Dissected: A Review of Current Knowledge and Analysis. *British J Plastic Surg.* 1996;49:143-149.

Young MA. Observer Agreement: Cumulative Effects of Repeated Ratings of the Same Samples and of Knowledge of Group Results. *J Speech Hear Res.* 1969;12:135-143.

Zajac DJ, Linville R. Voice Perturbations of Children with Perceived Hypernasality and Hoarseness. *Cleft Palate J.* 1989;7:696 - 705.

Zajac DJ, Mayo R. Aerodynamic and Temporal Aspects of Velopharyngeal Function in Normal Speakers. *J Speech Hear Res.* 1996;39:1199-1207.

Zajac DJ, Mayo R, Kataoka R. Developmental Aspects of Velopharyngeal Function: Evidence from Speech Aerodynamic Studies. *Paper presented at American Cleft palate-Craniofacial Assoc. Convention. New Orleans.* 1997.

Zajac DJ, Mayo R, Kataoka R, Kuo JY. Aerodynamic and Acoustic Characteristics of a Speaker with Turbulent Nasal Emission: A Case Report. *Cleft Palate-Craniofac J.* 1996;33:440-444.