Cohort 2012

MSc Speech and Language Therapy

Hypernasality and Cleft Palate Speech;
A comparison of inter-rater reliability using The Temple Street Scale and a Visual Analogue Scale,
In both trained and untrained Listeners.

Thesis by Sarah Jane Osborne (12011983).

Supervisor: Adj. Prof. Triona Sweeney

Date of Submission: 14/05/2014

Word Count: 7,088

Journal for Submission: Cleft Palate Craniofacial Journal
Abstract

**Background:** Perceptual listening is considered the ‘gold standard’ in the assessment of speech disorders relating to cleft palate (Kuehn and Molle 2000) but controversy remains as to the most appropriate type of rating scale to use. A descriptive scale for assessing nasality, The Temple Street Scale (TSS), has been tested for validity, reliability and acceptability (Sweeney 2000; Sweeney and Sell, 2008; Sweeney and Fennel 2009) and was subsequently adopted in the Cleft Audit Protocol for Speech–Augmented (John et al 2006). Little research exists for the use of a Visual Analogue Scale but Whitehill (2010) reported good reliability using this scale.

**Objectives:** To compare inter-rater reliability of hypernasality for 10 patients using The Temple Street Scale and a Visual Analogue scale and also to compare inter-rater reliability for both trained and untrained listeners.

**Methods:** Participants (7 trained and 20 untrained) were asked to rate 10 speech samples of varying degrees of hypernasality on two separate occasions. Percentage agreement, Mean Kappa Scores and Intraclass Correlation Coefficients were used to compare the reliability of the two scales and to compare the trained and untrained listener.

**Results:** Examination of percentage agreement showed slightly better agreement for the VAS compared to the TSS, though this was not the case when Mean Kappa was examined. Intraclass Correlation Coefficient showed very good results (> 0.81) for both scales and also showed slightly higher agreement for VAS, in line with Percentage agreement. The untrained listener scored better on both scales with Intrarater results of .984 and .98 for both the VAS and TSS respectfully compared to .937 and .950.

**Conclusions:** No significant difference between the two scales was observed. Both tools showed good reliability with percentage agreement and Intraclass Correlation Coefficients. The ease of use of the VAS may make it a good option for further investigation but interpretation of results for other multi-disciplinary team members difficult. Further research is needed to clarify.

**Keywords:** Hypernasality, The Temple Street Scale, Visual Analogue Scale.
Acknowledgements

Thanks to my project supervisor Adj. Professor Triona Sweeney for her help and expert guidance in this project. Thanks to Dr. Carol-Anne Murphy and Professor Sue Franklin for their help. Thanks to Brian and Selina for their support and encouragement and to my project partner Martina Dwyer. Sincere thanks to all of the participants who took part in this project and gave their time so generously.
Literature Review

Introduction:
While perceptual listening is now considered to be the “gold standard” in assessing speech affected by Cleft Palate and Velopharyngeal Dysfunction (Kuehn and Moller 2000), it has also been noted that confounding problems with this approach do exist (Sell et al 1999; Kent 1996; John et al 2006). For the perceptual assessment of nasality, one area of controversy is the type of rating scale used (Sell et al 1998). The Temple Street Scale (TSS) which uses a four point descriptive scale has been tested for validity and reliability with positive results (Sweeney 2000; Sweeney; Sell 2008; Sweeney and Fennell 2009). Parts of the TTS were subsequently adopted in the Cleft Audit Protocol for Speech-Augmented (John et al 2006). More recently, Whitehill (2010) has shown some promising results using a Visual Analogue Scale (VAS). This approach, although promising in the field of research, poses difficulties with regard to the clinical reporting of nasality and nasal airflow error as all members of the multidisciplinary team may not understand the ratings. The TSS is the predominantly used scale here in Ireland, however, various forms of assessment and scales exist globally. It is important that an international standardized system is achieved, whereby, protocols of speech assessment and measurement are established, including the use of a standardized rating scale that demonstrates validity, reliability and acceptability. This pilot study will examine if further research is warranted into the use of the VAS for hypernasality ratings by comparing it to the TSS which is an established and reliable tool. It will also examine inter-rater agreement for trained and untrained listeners as this is another area where further information is needed as the findings are unclear.

Background:
Children with a history of cleft palate are at risk for nasality and speech problems due to velopharyngeal dysfunction (VPD). These speech problems are characterised by anomalies of nasality, nasal airflow and articulation (Sweeney 2000). Nasality is defined as) the quality of voice determined by the balance of sound vibration in the oral, nasal and pharyngeal cavities during speech (Kummer 2002). Abnormal nasality can occur if there is obstruction in one of the cavities, causing hyponasality or cul-de sac nasality, or if there is velopharyngeal (VP)
dysfunction, causing hypernasality and or nasal emission (Sweeney 2009). Hypernasality is the excess nasal resonance perceived during the production of voiced oral sonorants, resulting from a coupling of the oral and nasal cavities, when the velopharyngeal valve is in the open position, (Sweeney 2009). This VP dysfunction can be due to velopharyngeal insufficiency (lack of sufficient tissue leads to ineffective VP closure), velopharyngeal incompetence (lack of neuromuscular competency in opening and closing the VP valve) or due to velopharyngeal mislearning (maladaptive articulation, no structural or neurological difficulties). Hypernasality is the hallmark perceptual deviant feature evident in the speech profiles of individuals with VP dysfunction (Dworkin et al 2004). It is estimated that 20-30% of cleft palate individuals suffer from hypernasality after primary closure (Riski 2005). As hypernasality is one of the hallmark features, it is clear to see that an international consensus and protocol would be beneficial. Sell (2005) reported that international consensus is still lacking, despite recommendations to establish clear protocols for speech assessment.

**Perceptual Assessment and Methodology Issues:**

Speech has long been recognised as one of the primary outcomes measures of palatal surgery (Grunwell et al 1993). The perceptual assessment of nasality constitutes an important aspect of a comprehensive assessment of speech for this client group (McWilliams et al 1990 and Sell et al 1999). Although as mentioned, perceptual listening is considered the ‘gold standard’ there are confounding problems with this approach (Sell et al 1999; Kent 1996; John et al 2006). Methodological issues around cleft speech include; consensus of terms, inconsistent and incomparable reporting measures, the type of speech elicited, the type of recording equipment used and crucially the type of rating scale (Sweeney and Sell 2008; Sell 2005; John et al 2006). The TSS has addressed these issues by clearly defining parameters of nasality, assessing various areas of speech shown to be essential, using high quality audio and video equipment, testing the assessment for validly and reliability, and using a rating scale with definitions of scalar terms (Peterson-Falzone et al 2001, Sweeney 2011). The TSS has also adopted speech samples, articulation and oral examinations from sections of the Great Ormond Street Speech Assessment (GOS.SP.ASS Sell et al 1994; 1999), which is the accepted national procedure for speech assessment in patients with cleft palate in the United Kingdom (Lohmander and Olsson 2003 p.65). However, unlike the TSS, the Great Ormond Street Speech

Little research exists on the use of VAS for hypernasality, Whitehill (2010) presented data from an unpublished dissertation by Cheng (2006) which showed good inter and intra-rater reliability for the VAS and the Direct Magnitude Estimation (DME). The VAS was reported to have greater reliability both for inter and intra rater-reliability and was described as ‘easier to use’ (Cheng 2006 p.23). It was also less timely in the calculating of results. The Sample size was small as 20 trained listeners assessed only 20 samples. VA scaling has proved to be a more accurate type of rating scales in other areas of speech assessment (Cannito et al 1997; Cannito et al 2005; Whitehill et al 2002; Whitehill et al 2007). A difficulty with using a VAS is that it may prove difficult to use in a multidisciplinary team setting where understanding and interpretation of results is important for all team members. As little research exists in this area, this pilot study aims to assess if more extensive research is warranted. The assessment used to elicit the speech sample was the TSS, the only variables will be the rating scale and the only area to be assessed is that of hypernasality (although the TSS outlines many more parameters associated with nasality and nasal airflow difficulties).

**The Rating Scale:**

The degree of hypernasality can be assessed using different rating scales which rely on listener perception. This listener perception means that results are open to interpretation. The choice of rating scale used can affect the listener’s ability to discern reliable differences along a dimension (Kent 1996). The appropriateness of a scaling method depends on ease of use, what is being measured (i.e. qualitative or quantitative) and whether the stimuli (hypernasality) is considered prosthetic or metathetic in nature (Stevens 1974; Stevens 1975). Descriptive category judgements and equal-appearing interval scales have been the preferred method of speech perception (McWilliams and Philips 1990; Sell 2005; Sweeney 2008).

Whitehill (2010) highlighted the import role that the rating scale plays in assessing nasality and the need for further research. She also stated that the most appropriate type of rating scale used may not be the same for clinical and research practice. This leads to a fundamental problem when comparing data for perceptual listening of hypernasality as a range of scales
are used. Descriptive scales, equal appearing interval scales and direct magnitude estimation are all present in the literature.

The Temple Street descriptive scale, for hypernasality is assessed as an individual parameter and was subsequently adopted by the Cleft Audit Protocol for Speech-Augmented (John et al 2006). It uses a 5 point descriptive scale with descriptions marked from a-e. Hypernasality is a perceptual phenomenon that eludes a single reference point that can be designated as normal, this is one of the reasons why a descriptive scale, as is used in the Temple Street scale, appears to be beneficial.

Whitehill (2010) has argued that nasality judgements are best rated using a continuous scale such as VAS or Direct Magnitude Estimation (DME). In the VAS (Flynn et al 2004) severity is calculated by distance from 0, (Cannito et al 1997), where 0 represents normal nasality and the 10mm end point marks severe hypernasality. Hypernasality is a prosthetic signal and is qualitative and substitutive in nature and by using a continuous line, it is suggested that more accurate results are obtained (Redenbaugh and Reich 1985; Whitehill et al 2002). There is evidence showing that visual analogue scales have ‘superior metrical characteristics than discrete scales’, meaning a wider range of statistical methods can be applied to VAS. However, one suggested reason that researchers have steered away from VAS (as well as DMS), is that interpretation of these results are time consuming and impractical (Reps and Funk 2008).

Traditionally, judgements have been made using the equal appearing interval (EAI) scales EAI (Morris, Shelton and McWilliams 1973). Different types of EAI scales exist from 3-11 point scales and an area of difficulty has been the manner in which each parameter and scalar point is defined (Henningsson et al 2008). Another disadvantages of EAI scaling is that listeners tend not to treat the points on the scale as equal intervals (Edwin and Chi 2004; Whitehill 2002) and the points are often not well defined (Sweeney 1996). Although traditionally preferred, EAI scale research has been flawed regarding the measures used to assess reliability. Interval rating scales are thought not to be the most appropriate scale when assessing nasality (Zraick and Liss 2000; Whitehill 2002).

In hypernasality there is an increased chance of co-existing disorders in cleft palate speech, making hypernasality rating difficult when the parameter is not assessed individually but
solely under the parameter ‘nasality’ (Kent 1996; Kataoka et al 2001; Imatoni 2005). It is also multidimensional, and other factors such as rate, pitch and loudness effect ratings, so EAI and VAS scales that rate nasality together have not been considered for discussion. Rating scales used must prove to be robust, it is important that studies correctly measure reliability to allow for better comparison between studies in the future. It is important that reliable outcome measures are reported for surgical, prosthetic and speech therapy. Outcome measures are also needed for clinical audit, treatment efficacy, and research and cross-centre studies. They are an important part of the move towards evidence based practice. McCauley (1989 p.29) further emphasised this when he stated “the quality of clinical action depends upon the quality of measure” and the correct choice of rating scale is key.

The trained versus the untrained listener:

A further variable which may influence reliability of perceptual assessment is listener experience. Early studies which examined nasality indicated that the role of the experienced listener may not be as large as expected (Bradford et al 1964), and that trained listeners are affected by areas such as bias (Kent 1996), information obtained in the case history, physical features (Lopez, 1989) and familiarity with the speaker (Wilcock et al 1990; Tjaden; Liss 1995). Another variable is different internalized standards for ‘normal’ nasality (Whitehill 2010). The use of training and anchors (external standards) in the evaluation process have been suggested as possible ways to minimise bias (Chan and Yiu 2002, Gerratt et al 1993; Kreiman et al 1993). Some more recent studies have shown correlations between inexperienced raters were generally lower than experienced raters for nasality difficulties (Hayden and Klimacka 2000). This is further supported by Keuning et al (1999) who found that experienced listeners had higher reliability than inexperienced listeners while examine hypernasality, nasal emissions, intelligibility and misarticulations. Sell et al (2005) noted that no training was provided by the researchers and it was unclear what trained listeners meant. While much research exists around trained versus untrained listeners it is unclear when assessing nasality how much it affects reliability, studies show contradictory and varying degrees of results and trained and untrained listeners are rarely defined. In a critical review of the methodology used for assessing speech in cleft palate, Lohmander and Olson (2004 pg.64) stated that their most ‘distressing’ finding was the lack of information on reliability and only approximately half of the 88 articles reviewed included reliability measures. Experience is valuable for many
reasons but it does not seem to be the ‘prima facie evidence of inter-judge agreement’ (Kent 1996 pg.17) and Lohmander argued that ratings should be carried out by ‘naïve listeners’.

As outlined, a number of variables are known to affect the perceptual listening of hypernasal speech, Whitehill (2010) reported five parameters, two of which will be examined here. The first being the type of rating scale used and the second being the effect that listener training has on reliability. Other factors include the speech material used, listener variables (such as the first language of the speaker) as well as speaker variables (such as the use of headphones versus speakers).
Aims

Aim 1: to compare inter-rater reliability of hypernasality for 10 patients using the Temple Street Scale and a Visual Analogue Scale.

Aim 2: to compare inter-rater reliability on both scales for untrained and trained listeners.
Method

Speech samples

Speech Samples to be rated were obtained from the Speech and Language Therapy video archive of children who have attended The Children’s University Hospital for cleft related speech problems (see appendix 1 for medical diagnosis). All samples had been previously recorded using high quality audio/video digital recorder Zoom Q3 HD. 10 videos were selected that ranged in severity from normal to severely hypernasal. Samples covered an age range of 5-19 years and contained an equal number of male and female samples (appendix 1). All speakers were Irish and had English as their first language. (One speaker had Polish parents but English was his first language). Researchers rated speakers and cross checked with previous ratings given by Adjunct Professor Triona Sweeney so that a range of severity could be selected, controls were also included (appendix 1). Samples that showed inconsistency were not included. It was not possible to get 10 pure samples, where hypernasality was the only parameter affecting speech, and as such samples were chosen where hypernasality was the predominate speech difficulty with other factors affecting intelligibility kept to a minimum. All speakers were Irish and were from various parts of the country (appendix 1). Previous ratings had been done using The Temple Street Scale and as such recordings contained examples of rote speech, conversation speech and The Test Sentences of The Temple Street Scale (appendix 2).

Participants

(All participants were Irish with English as their first language)

Trained; 16 speech and language therapist were contacted from a Cleft Special Interest Group. Training involved a 1 day course (see appendix 3 for typical course outline), which included a definition of terms of all parameters assessed in the TSS (see appendix 4 for hypernasality definition), an audio demonstration of each and practice rating 10-12 cases which were hypernasal. Of those contacted, 7 signed consent forms and completed both parts of the project, unfortunately due to time restrictions and unforeseen events 1 dropped out before the project started and 3 listeners who had completed part 1 had to withdraw
from the project. Of those that completed the project, 5 worked in cleft specialist centres in Dublin and one worked in the community (mid-lands).

**Untrained:** 20 Speech and Language Therapy Students from the University of Limerick MSc took part. Students were 2nd year master students who were 3 months away from qualifying. None had received TSS training and exposure to hypernasality was minimal.

**Viewing and Rating of samples**

1) **Trained listeners; rating of samples methodology**

The project was designed so that Clinicians could carry out speech sample ratings in their own time, while still allowing it to be counted towards their Continual Professional Development (CPD) hours. It was hoped that by doing this, the number of participants would be higher than if asked to attend two separate training days.

E-mails were sent out to members of the Cleft SIG group with a brief introduction of the researchers and the project. A more detailed information word document outlining the project was included in an attachment along with consent forms (appendix 5 and 6). Anyone interested in partaking in the project was asked to sign and return the consent form via post or e-mail. Once consent forms were received, letters were sent out with instructions for completing part 1 of the project (appendix 7). This letter also included the password for the key. Keys were subsequently sent out with instructions on opening (appendix 5). This was done via registered post to ensure privacy. Sending out the letter prior to the key allowed for the separate arrival of the password (this was included in the letter with part 1 instructions). This was done to further ensure the privacy of the data. Participants were asked to destroy the password once received.

The speech samples chosen (as previously outlined) were anonymised and transferred onto encrypted keys in random order. Keys were posted out to the trained listeners via registered post. Participants were asked to watch the video’s using high quality speakers or headphones such as Sennheiser 90 headphones or Genelec speakers.

**Part 1; the TSS**
Participants were asked to complete a demographics questionnaire (appendix 9a) before being asked to complete Part 1 of the project. This survey was carried out on Survey Monkey, a web survey development cloud based company. As part of the demographics, they were asked to advise which speakers they intended to use. 3 participants marked the Sennheiser 90 headphones and 3 marked the Genelec speakers and 1 participants did not respond to this question. Other questions on the survey were related to demographics and can be seen in appendix 9b). Part 1 of the project involved rating each video using The Temple Street Scale for hypernasality (appendix 10, for the full TSS please see appendix 11) to rate the speech samples. The only parameter measured was that of hypernasality. The survey was carried out online via Survey Monkey. To do this the TSS of Nasality had to be incorporated onto a questionnaire sheet on the website (appendix 7). Consistency was not measured and any inconsistent samples were omitted therefore these parameters which are normally assessed under ‘Nasality’ in the TSS were dropped from the online form. Online data collection was done to allow ease of use for participants and data collection for researchers.

Part 2; the VAS

To create the VAS for Part 2, a 10cm line was drawn using Microsoft word, the left end indicated normal resonance and the right end indicated severely hypernasal. Instructions on part 2 were sent out via e-mail and the VAS form was attached (appendix 11 and 12). Participants were asked to watch the same videos again in the same order and rate each by placing a vertical line along the horizontal line on the form that matched the perceived magnitude. Participants were asked to complete Part 2 four weeks later to minimise memory effects and therefore avoid previous ratings influencing part 2. Compliance for completion of part 2 was 27-37 days later. The results for the VAS were collected via post or e-mail. Unfortunate due to cost restrictions we were unable to carry out the rating for the VAS online, as was done in Part 1.

2) Untrained listeners; ratings of samples methodology

The untrained participants were asked to attend two video sessions which were separated by 4 weeks. Each session was 1 hour long, this was the same timeframe that the trained participants were asked to allocate. Videos were presented on campus in a small lecture theatre using a high quality projector system with high quality speakers (NEC NP-P401W Entry
Level Installation 4000-Lumens LCD Projector with 16W Speakers) and played through VCL media player. Videos were presented in the same order as those in the trained group and the TSS was completed first as with the trained group. An introduction of the project was given and students who agreed to take part were asked to sign the consent forms. Students were given a chance to read the descriptive scales before videos were played. No further information was given and students were simply asked to rate each video according to what rating they deemed most appropriate. Ratings were collected online via Survey Monkey for Part 1 as was the case with the trained listeners. For Part 2, students were asked to mark the VAS form with one vertical line at the most appropriate point. Results were collected by hand by researchers.

**Data Analysis**

To examine reliability results were collected and inputted into SPSS 10, for Windows. Six groups of data were obtained including; results from the trained, untrained and a combination of the two (total) for both the TSS and the VAS. To analyse the VAS, the results of each of the ratings were measured using a ruler and given a measure to the nearest decimal point. To assess the rate of agreement on the VAS, 5 points of agreement were given, so that a result could match anything within a +/- 1cm range. To analyse and compare the two scales (TSS and VAS) both non-parametric and parametric testing was carried out. This included percentage agreement, mean Kappa scores and Intra-Rater Correlation Coefficients. In analysing mean Kappa and percentage agreement scores, paired scores were obtained for each rater and a subsequent mean value was calculated for each group. Data was tested for normality using The Sharpio-Wilk Test (appropriate for smaller sample sizes). Data was shown to be normally distrusted (p>0.05) allowing for Intraclass Correlation. ICC’s to be carried out were; Two Way Mixed absolute for measuring absolute agreement and Two Way Consistency.

**Ethical Issues and destruction of data**

All listeners gave fully informed consent for the project. It was made clear that no identifying information would be used. The speech samples which had previously been recorded, had all been done with informed consent for speech records, teaching and research. As clinicians were to rate the samples in their own time confidentiality of patient video files was of upmost importance. Fully encrypted keys were used as oppose to simply password protected, to
ensure privacy was at the highest level. Passwords for these keys were also sent separately to ensure that only the participants could open them. All samples were posted via registered post. Speech samples themselves did not contain any identifying information and samples were simply coded 1-10. Participants were asked to delete all files from the keys and return them via registered post once both parts of the project were completed. This was done so that deletion of all files could be ensured in line with our ethics outline to The Children’s University Hospital, Dublin. Any participants who preferred could also arrange to have the key collected. All hard copies of data including consent forms and the VAS forms will be kept in a locked filing cabinet in the Children’s University Hospital for 3 years, after which they will be destroyed. All digital data files including Microsoft Excel Sheets, SPSS sheets, coded information of speech samples and participants will be stored on an encrypted key as above and will be deleted in 5 years.
Results

Agreement and Kappa scores are outlined in the tables below for all 27 participants. The mean of the paired scores for each rater as well as the range from highest to lowest is included.

Table 1; Percentage Agreement

<table>
<thead>
<tr>
<th>% agreement</th>
<th>The Temple Street Scale (TSS)</th>
<th>Range</th>
<th>Visual Analogue Scale (VAS)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total:</td>
<td>38.2%</td>
<td>10-80</td>
<td>40.3%</td>
<td>10-80</td>
</tr>
<tr>
<td>Trained:</td>
<td>36.19%</td>
<td>20-70</td>
<td>39.5%</td>
<td>20-80</td>
</tr>
<tr>
<td>Untrained</td>
<td>39%</td>
<td>10-80</td>
<td>43.7%</td>
<td>10-80</td>
</tr>
</tbody>
</table>

Table 2; Mean Kappa Scores

<table>
<thead>
<tr>
<th>Mean Kappa Score</th>
<th>The Temple Street Scale (TSS)</th>
<th>Range</th>
<th>Visual Analogue Scale (VAS)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total:</td>
<td>0.25</td>
<td>-0.18 - 0.88</td>
<td>0.14</td>
<td>-0.031 - 0.92</td>
</tr>
<tr>
<td>Trained:</td>
<td>0.23</td>
<td>-0.34 - 0.40</td>
<td>0.19</td>
<td>-0.031 - 0.84</td>
</tr>
<tr>
<td>Untrained:</td>
<td>0.27</td>
<td>-0.18 - 0.88</td>
<td>0.21</td>
<td>-0.20 - 0.92</td>
</tr>
</tbody>
</table>

Table 3; Interpretation of KAPPA statistic (K) and inter-class correlation coefficients (ICC) as outlined by Landis and Loch 1997)

<table>
<thead>
<tr>
<th>K</th>
<th>Strength of Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00- 0.20</td>
<td>Poor</td>
</tr>
<tr>
<td>0.21–0.40</td>
<td>Slight</td>
</tr>
<tr>
<td>0.41–0.60</td>
<td>Fair</td>
</tr>
<tr>
<td>0.61–0.80</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.81–1.00</td>
<td>Good</td>
</tr>
<tr>
<td>1.01–1.99</td>
<td>Very good</td>
</tr>
</tbody>
</table>

Table 1 lists results for the Percentage Agreement and also includes the ranges. Results show the VAS scored more favourable, demonstrating a higher percentage agreement for all 3
groups (trained, untrained and total). The untrained listener was more reliable than the trained on the VAS and the TSS. Results were reflective of the wide range that was seen and no major outliers were noted. Results for the two scales were seen to be comparable with results of 39.5%, 43.7% and 40.3% for the VAS and 36.19%, 39% and 38.2% for the TSS.

**Table 2** Mean Kappa Scores produced were not in line with those found via percentage agreement. Results for the VAS were seen to be lower in all three groups (trained, untrained, and the group as a whole). The TSS scored more favourable showing fair agreement compared to the VAS which only showed slight agreement (Landis and Loch 1997). In interpreting Kappa the closer one gets to a coding of 1, the higher the rate of agreement, with 1 demonstrating perfect agreement and 0 when there is no agreement other than what would be expected by chance (Haney et al., 1998). Although TSS scored greater reliability (fair as oppose to slight), data obtained was noted again to be similar between the two scales. When comparing the trained to untrained listener the untrained listener scored slightly higher again (0.27 as oppose to 0.23) for the TSS and also for the VAS (0.21 as oppose to .19).

**Table 3** was included for easy reference and interpretation of results from Table 2. (Mean Kappa) and can also be used to interpret figures 1 and 2 (ICC’s)

**A test of normality:** The Sharpio-Wilk test of normality showed data to be normally distributed. Significance range was between 0.14-0.705 for the VAS and 0.8-0.89 for the TSS and thus the null hypothesis was accepted and the ICC was deemed an appropriate measure of agreement.

**Figure 1 (a); Intraclass Correlation Coefficient Analysis for absolute agreement.**
Results of the ICC were very good according to Altman (1991). ICC for absolute agreement measures how much the raters agree on the values that they assign. Figure (a) contains data for all 3 groups (trained, untrained and total) and is laid out for ease of interpretation of aim 1), comparing TSS and VAS. It is clear from the visual representation that although the results are similar VAS scored higher demonstrating greater reliability for the trained listener. The group as a whole also scored higher on the VAS. Untrained clinicians were again seen to show great reliability on both scales.

Figure 1 b) allows for a clear representation of aim 2) the trained versus the untrained listener. From the figure it is clear that the untrained listener scored higher showing greater absolute agreement in both the TSS and VAS.
While figure 1 answers our research question Figure 2 allows us to examine whether participants were consistent in the order in which they rated the videos and allows another dimension of reliability to be added. This adds further weight to the reliability of the results; according to Altman (1991) reliability was very good.
Discussion

The primary aim of this study was to compare the TSS and VAS in the perceptual evaluation of hypernasality and the second was to compare the trained and untrained listener. To do this, Percentage agreement, Kappa scores and ICC values were calculated to allow inter rater agreement to be analysed. Using the three different types of analysis allowed a clearer picture of agreement between the two scales and also allowed more comparisons to the literature. The VAS was seen to demonstrate very good reliability and the untrained listener showed a higher inter-rater agreement for both scales. The untrained group also showed higher agreement.

Percentage agreement was the most common method used for analysing hypernasality in the literature. It has the unique advantage of putting more complicated statistics into context, allowing easy understanding of results at a practical level (Agresti 1996) and therefore was a good place to start. Results were similar between the two scales with 38.2% for the total TSS and 40.3% for the total VAS obtained. This demonstrated a good outcome for VAS and good potential for future use in assessing hypernasality. This trend was seen in all three groups where the VAS scored 39.5% for the trained listener, this was compared to 36.19% in the TSS and 43.7% for the untrained compared to 39% in the TSS. The range on both scales was large with 10-80% on both, this affected the result lowering percentage agreement below the predicted.

Reliability studies on the use of the temple street scale reported much higher percentage agreements and results showed a range from 64-80%. Inter-rater reliability studies that have used 5 point descriptor scales as oppose to 4 in assessing the nasality demonstrated lower reliability than those of the TSS. As the type of rating scale used was the same (descriptive), percentage agreements results have been included for reference but should not be directly compared. When examining the TSS, Sell et al (2009) obtained 88.9%, Lohnmander-Agerson (1999) obtained 80%, and Lohmander et al (2002), obtained ranges from 64-69% for hypernasality all of which were carried out on trained listeners. Sell et al (1994) obtained a value of 84-92% agreement, though ratings for this were obtained on a 3 point scale and so one would expect percentage agreement results to be slightly higher as percentage
agreement is not adjusted to chance. As expected, lower results were then seen when 5 point descriptive scale was used. Results by Lideman et al (2005), Person et al (2006) and Lohmander et al (2008) were lower at 39%, 44-55% and 39% respectfully. These results were a lot more comparable to ours with our trained listeners scoring a 36.19% on the TSS and 39.5% on the VAS. Although these results were similar to ours, as mentioned these ratings of percentage agreement were made on a 5 point scale not 4 point scale and so cannot be directly compared.

Results on the two scales were similar as was expected, this is in line with other areas of speech (and voice) assessments where VAS is more commonly used. In assessing hypernasality for this project no other parameters were measured. Factors such as voice quality, loudness and pitch are thought to effect perception of nasality (Zraick 1999). Hypernasality is considered to be multidimensional (Zraick 1999) and it is possible that isolating the parameter of nasality on its own may have lowered results.

No research was found (using PubMed and Google Scholar as such parameters) for the use of VAS in hypernasality. The research question was posed from a presentation from Whitehill (2010) which reported positive findings from an unpublished dissertation Cheng (2006) who reported ICC’s.

As mentioned, one of the limitations of percentage agreement is that it does not take into account chance agreement. This is particular pertinent when the variability of observations are small. Percentage agreement also has a potentially infinite range and therefore reported values cannot be interpreted within a standard range.

Kappa Scores were also carried out to look at inter rater agreement. Cohen’s kappa coefficient allows for a statistical measure of qualitative items to be carried out. As it takes into account chance agreement, it is generally thought to be a more robust measure than simple percent agreement calculation. Kappa is computed as:

\[
\kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}
\]

Where

\( Pr(a) = \) proportion of units on which the raters agree

\( Pr(e) = \) the proportion of units for which agreement is expected by chance
Cohen's kappa coefficient only measures agreement between two raters and as such a mean was ascertained. Kreiman et al (1993) reported that averaging data across clinicians can weaken aspects of the individual ratings. Paired agreements were calculated as per Cohen (1960) and the mean values obtained.

All values were lower than expected. The TSS scored more favourably showing fair agreement compared to the VAS which only showed slight agreement (Landis and Loch 1997). This was lower than previous research including a range of .4-.6 (Sweeney 2000) and .8 (CSAG 1998) for trained listeners and which were far more favorable than ours at 0.23-0.27. This may have been due to the multifactorial nature of hypernasality as only one parameter of speech was measured. Only slight agreement was seen on the VAS.

A number of negative values were obtained when calculating paired scores for Kappa on the VAS and the range was much larger at -0.031-0.92 compared to a range of -0.34 – 0.88 for the TSS. A negative kappa value implies that raters agreed less than would be considered by chance (Tinsely 1975). Tinsley (1975) stated that Kappa scores can over correct for chance agreement and therefore can be overly conservative. There were a higher number of negative values obtained on the VAS compared to the TSS. The results when looking at kappa scores alone would suggest the TSS is the more reliable tool for assessing hypernasality.

Results for mean kappa did not appear to be in line with those calculated by simple percentage agreement and higher values were obtained for the TSS (where higher VAS agreement was noted using simple percentage agreement). A reason for this may have been that according to Shrout and Fleiss (1979), Kappa agreement measure are ideal when examining inter-rater reliability when the absolute value of the rating matters. It is there for an ideal method to use with the TSS as this is a 4 point scale. ICC’s on the other hand are deemed to be a more appropriate means of analysis for continuous scales such as the VAS as it measures the reliability ratio directly (Shrout and Fleiss 1979). This would account for the VAS scoring lower than the TSS.

The Sharpio-Wilk test which is a standard test of normality was used to ascertain if the data was normally distributed and therefore parametric. It is the preferred test when sample sizes are small (Bartko 1976; Garson 2012). It uses the null hypothesis principle to check whether a sample came from a normally distributed population (Sharpio and Wilk 1965) and is
recommended before carrying out ICC’s (Bartko 1976). A p-value > 0.05 means the null hypothesis is accepted and therefore the data is ‘normal’ on the other hand a p-value < 0.05 means that the null hypothesis is rejected and therefore the distribution is not normal. As p ranged from 0.14-0.89 data was considered normal.

**Intrarater Reliability Correlation Coefficients** are a good measure to test agreement on continuous scales (Bartko 1976; Shout et al 1976) and as such these were also carried out. ICC’s operate on data structured as groups, rather than data structured as paired observations and is advantageous for looking at consensus between groups for parametric data. Data was analyzed using the Two-Way mixed model first, looking at absolute agreement, followed by consistency. This was done by using Analysis of variance (ANOVA), f tests, and with a 95% confidence interval in SPSS as was outlined by McGraw (1996). Results for the ICC were good (Altman 1991) and were in line with those found on simple percentage agreement which showed that inter rater reliability on the VAS was better. This was also the case for the untrained listener. Results for percentage agreement, kappa mean and ICC’s need to be interpreted with caution, Kreiman et al 1993 caution’s that averaging data across raters weakens the data.

The second aim of the study was to evaluate listener experience on reliability. Results all demonstrated the untrained listener to be preferable. The literature is much divided on this and how much listener experience plays a role. Our results fall in line with early research that stated that experienced raters tend to be less reliable as they are affected by bias due to internalised standards. More recent studies have seen the trained rater show greater inter-rater reliability. The degree of nasality assigned also appears to be a big factor for rating when comparing the two groups (Fletcher 1976), with native raters tending to score towards the more severe end of the spectrum and showing greater variability in their results. This was seen in the results obtained with an increased number at either end of the mean range for VAS, though agreement was still higher.
Difficulties in methodology

Listening conditions

A couple of factors that may have affected the results when comparing the trained and untrained listener are worth noting. First, the listening conditions between the two groups had some discrepancies that were unfortunately unavoidable. The untrained listeners viewed samples by attended ‘a training’ day where conditions were controlled, so that room acoustics, volume, video player and speakers were consistent for all listeners. As trained listeners viewed the speech samples individually and in their own time speaking conditions would not have been the same. Researchers did try to minimise this effect by outlining appropriate speakers to be used. All untrained listeners also listened to the speech samples at the same time, uninterrupted and without distraction. The untrained listener also sat in close proximity to other raters which may have influenced results, though researchers noted that they did not appear to watch their peers and all surveys were kept anonymous.

Difficulty with measuring ordinal and scalar data

It is also noteworthy that there is some difficult with selecting the most appropriate method of statistics as the two scales differed. Kappa is the preferred statistic for the estimation of interobserver agreement for nominal and ordinal data (Haley and Osberg 1989) and so is a good measure in analysing inter-rater reliability for TSS but possibly not so reliable for VAS where data is scalar. ICC’s have been suggested to be more appropriate measures of agreement with scalar data (Bartko1976). Another difficulty noted by the researchers was in defining the cut of points in analysing agreement for VAS, research using the VAS was limited and it was unclear from the literature how parameters were interpreted. The VAS clearly scored better when using percentage agreement and ICC results however this was not the case with kappa. These results suggested that the TSS is the more reliable tool, it was unclear to the researchers as to the reason for this result discrepancy.
Presentation of data

The VAS for all 10 videos was presented on a single sheet. This meant that previous ratings may have influenced subsequent ones on the scale. This was not the case for the TSS, once a score was rated they were unable to go back and view the rating.
Conclusion

Both scales demonstrated very good reliability when looking at percentage agreement and ICC results. This suggests that the VAS is a reliable tool for the assessment of hypernasality and does warrant further investigation. Both scales were seen to show similar results, both very good in the ICC analysis (Altman 1991). However, one difficulty with the use of a VAS may be with the interpretation of the results by other team members when working within an MDT environment. Whitehill (2010) notes that the most appropriate tool for rating may not be the same for clinical and research purposes. Our findings were positive for the use of VAS in the rating of hypernasality and therefore we would recommend that more research is needed with its use in assessing hypernasality. More specifically, research with larger participant’s numbers and qualitative analysis of ease of use within an MDT setting. The low kappa scores highlight the fact that although perceptual listening is considered the ‘gold standard’ of speech assessment many problems still remain. This paper aimed to examine just two of these variables; the rating scale and that of listener experience. Listener experience was seen to have a negative effect on reliability, this is in line with earlier research (Kent 1994). Listener experience is thought to affect bias, with previous exposure creating internal standards. The literature remains divided on this debate and this paper certainly contributes to the bias arguments.

It important that research continues in this area so that the most reliable and valid scales are chosen both for clinical use and in research practices. A universal standard needs to be obtained so that comparisons of data can be easily made. In the move towards best evidence based practice it is essential that all areas continue to be re-assessed and new findings are constantly compared to current practices so that the most reliable tools can be used both for clinical evaluation and in the areas of research.
References


Appendices

Appendix 1; Speech Samples

<table>
<thead>
<tr>
<th>Video No.</th>
<th>Assigned Rating – Prof., Triona Sweeney.</th>
<th>Gender</th>
<th>Age</th>
<th>Area</th>
<th>Medical Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>d</td>
<td>F</td>
<td>5.03</td>
<td>Kerry</td>
<td>*UCLP</td>
</tr>
<tr>
<td>2</td>
<td>WNL</td>
<td>F</td>
<td>5</td>
<td>Louth</td>
<td>*CP</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>M</td>
<td>5.09</td>
<td>West Meath</td>
<td>CP</td>
</tr>
<tr>
<td>4</td>
<td>a</td>
<td>M</td>
<td>7.11</td>
<td>Kerry</td>
<td>CP</td>
</tr>
<tr>
<td>5</td>
<td>d/e</td>
<td>F</td>
<td>5.08</td>
<td>Dublin</td>
<td>CP</td>
</tr>
<tr>
<td>6</td>
<td>E</td>
<td>M</td>
<td>19.08</td>
<td>Limerick</td>
<td>*VPI</td>
</tr>
<tr>
<td>7</td>
<td>C</td>
<td>M</td>
<td>5.09</td>
<td>Dublin</td>
<td>*UCLP</td>
</tr>
<tr>
<td>8</td>
<td>WNL</td>
<td>M</td>
<td>10.03</td>
<td>Cork</td>
<td>*BCLP</td>
</tr>
<tr>
<td>9</td>
<td>b</td>
<td>F</td>
<td>9.08</td>
<td>West Meath</td>
<td>*SMCP</td>
</tr>
<tr>
<td>10</td>
<td>b</td>
<td>F</td>
<td>11.05</td>
<td>Dublin</td>
<td>SMCP</td>
</tr>
</tbody>
</table>

*CP; cleft palate, VPI; Velopharyngeal insufficiency, UCLP Infants with complete unilateral cleft lip and palate, BCLP; children with bilateral cleft lip and palate, SMCP; Submucous cleft palate

Appendix 2; The Temple Street Scale (Sweeney 2000), Speech samples elicited.

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. Vicky’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

32
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 3: TSS course outline.** (courtesy of Adjunct Professor Triona Sweeney).

**Temple Street Scale Training Workshop**

9.30 - 11
- Perceptual assessment
- Problems with Perceptual assessment
- Speech samples
- **Definition of Terms** (See Appendix 2)
- Detailed Descriptive Scale-

11.30 -12
- Reliability & Validity
- Assessment Booklet Slide
- Audio/video

12 – 12.30  **Group Listening- Audio/Video examples**
- Hypernasal example 1 - rating b
- Hypernasal example 2 - rating a2
- Hypernasal example 3 - rating c2
- Hypernasal example 4 – rating e
- Hyponasal example 1
- Hyponasal example 2

1.30 – 2.30  **TSS audio/Video examples**
- Cull de Sac Hyper C
- Nasal Emission Hyper A
- Nasal turbulence Hyper B
- ANF

3.00 – 4.30  **Reliability LISTENING - TSS Speech Samples**
(Examples given of each of the below)
Appendix 4: Definition of Hypernasality Sweeney (2000)

**HYPERNASALITY:** The occurrence of excessive nasal resonance perceived during speech production resulting from the coupling of oral and nasal resonating cavities, where there is an increase in the ratio of nasopharyngeal port to oropharyngeal port size.
Appendix 5: Information sheet outlining our project

To Whom It May Concern:

We are MSc in Speech and Language Therapy students from the University of Limerick conducting our final year research project on the reliability of hypernasality ratings in general cleft assessments. This is a question that has been raised as a result of Cleft Audit Cleft Audit Protocol for Speech – Augmented (CAPS-A) and Americleft training. We are now eagerly looking for SLTs to participate in the study. The project will be carried out under the supervision of Adj. Professor Triona Sweeney, Clinical Specialist/Adjunct Professor at Temple Street Children’s University Hospital.

Participation is completely voluntary, and you may withdraw your participation from the study at any time without consequence. We appreciate that you are all very busy but the benefits of participation in this study far outweigh any negative consequences. Your input would be of great value to this project and the process can be included in your yearly CPD profile for IASLT.

Although it is now acknowledged that the perceptual assessment is the gold standard for assessment of speech disorders related to cleft palate and velopharyngeal dysfunction (Kuehn and Moller 2000), it is also acknowledged there are confounding problems with this approach (Sell et al. 1999, Kent 1996, John et al. 2006). For the perceptual assessment of nasality one area of controversy is the type of rating scale used. More recently, Whitehill (2010) reported good reliability for a visual analogue scale (VAS), and although this approach shows promise for research, the VAS can make clinical reporting of nasality and nasal airflow errors difficult; all members of the multidisciplinary team may not understand the ratings. A descriptive scale for assessing nasality and nasal airflow, the Temple Street Scale (TSS), has been tested for validity, reliability and acceptability (Sweeney, 2000; Sweeney and Sell, 2008; Sweeney & Fennell, 2009). This approach assesses nasality and nasal airflow errors, using detailed definitions of terms and descriptive definitions of scalar points. The aims of this project will be:

1) To compare inter-rater reliability for assessment of hypernasality for 10 patients using the Temple Street Scale and a Visual Analogue Scale and

2) To compare inter-rater reliability on both scales for untrained and trained listeners

To carry out this project, you will be provided with an encrypted memory key containing 10 video samples. You will be asked to watch and listen to these samples on a computer and
To provide ratings of hypernasality only. You will be required to complete TSS ratings on one occasion and to then complete VAS ratings on the same 10 samples one month later. Both sets of ratings should take about one hour. You will then be required to post your ratings to Adjunct Professor Triona Sweeney at Temple Street Children’s University Hospital for analysis. The results will be analysed using Kappa scores to assess reliability. Mean percentage agreement scores for the above will also be calculated. Mean reliability scores will be compared and analysed for the two rating scales.

All information will remain entirely confidential and anonymous. To ensure confidentiality, all participants will be required to return encryption keys (via hand or registered post) at the end of the study so that their contents can be destroyed. Although information may be published in an article or presented at an academic conference in the future, nothing of a personal or distinguishing nature will be included.

If you agree to participate in this study, please sign the attached consent form and return via post or email by the 13th of January. Please see contact details below. Please do not hesitate to contact either of us if you have any query regarding the study. Many thanks for taking the time out to read this and we hope to hear from you soon.

Yours Sincerely,

Sarah Jane Osborne          Martina Dwyer          Prof. Triona Sweeney

Directors: Mr. Donal Walsh (Chairperson), Ms. Siobhan Brady, Dr. Liam Clafley, Ms. Suzanne Dempsey, Dr. Michael Drumm, Mr. John Fitzpatrick, Mr. Fionn MacCumhaill, Mr. Derek McGrath, Mr. Frank McManus, Sr. Margherita Rock, Mr. Sean Sheehan

36
Contact Details

Post:

Martina Dwyer c/o Geraldine Morrissey,
UL Speech and Language Therapy Clinic,
Clinical Therapies Department,
University of Limerick,
Castletroy,
Limerick.

Research Assistant 1: Sarah Jane Osborne
Email: sosbornes@gmail.com

Research Assistant 2: Martina Dwyer (MSc in Speech and Language Therapy student)
Email Address: martinadwyer4@gmail.com

Principal Investigator: Prof. Triona Sweeney
Email Addresses (please contact both):
Triona.Sweeney@cuh.ie
sweenyt@eircom.net

Company Registration N0.351404    Charity No. CHY 229
Directors: Mr. Donal Walsh (Chairperson), Ms. Siobhan Brady, Dr. Liam Clalley, Ms. Suzanne Dempsey, Dr. Michael Drumm,
Mr. John Fitzpatrick, Mr. Fionn MacCumhaill, Mr. Derek McGrath, Mr. Frank McManus, Sr. Margherita Rock, Mr. Sean
Sheehan
Appendix 6; Consent forms

Consent Form

Research Study: To evaluate and compare the assessment of hypernasality, in The Temple Street Scale of Nasality and Nasal Airflow and using a Visual Analogue Scale for rater reliability in both the trained and untrained listener.

- I have read and understand the attached information sheet, and by signing below, I consent to participate in this study.
- I understand the aims of the study
- I understand that all information obtained will be kept confidential
- I understand that I also have the right to change my mind about participating in the study for a short period after the study has concluded.

Date: _______________________________________________________________________

Signed: _______________________________________________________________________

Print Name (Block Capitals): _______________________________________________________________________

Company Registration N0.351404  Charity No. CHY 229Directors: Mr. Donal Walsh (Chairperson), Ms. Siobhan Brady, Dr. Liam Claffey, Ms. Suzanne Dempsey, Dr. Michael Drumm, Mr. John Fitzpatrick, Mr. Fionn MacCumhaill, Mr. Derek McGrath, Mr. Frank McManus, Sr. Margherita Rock, Mr. Sean
Appendix 7; Letter of Instruction for Part 1

11/02/2014

To all participants,

Thank you for your continued participation in our study on hypernasality. You will be receiving your encrypted key, with video files that will need to be rated in the next day or so. The password for opening the key is;

********

Please see accompanying letter for instructions on opening the video files.

This is Part 1 of a two part project. This email will outline instructions for completion of Part 1. Part 1 will include some demographic questions as well as the first set of hypernasality ratings (Temple St. Scale ratings). Details pertaining to Part 2 (i.e. second set of ratings via a visual analogue scale) will be imparted to participants in three weeks’ time. Please allow approximately one to one and a half hours to complete Part 1.

**Firstly** please follow the link below to complete some demographic questions on Survey Monkey, choosing the most appropriate answer in the drop down menu,

https://www.surveymonkey.com/s/7RNBTDD

(If clicking on this link does not lead you to the survey, try copying and pasting the link into the search bar on your internet browser).

**Next** you will be required to watch the videos and complete the Temple St. Scale ratings. Please ensure not to copy the video files to your own computer. You need to watch and listen to the videos in numerical order and rate each one immediately after. **You will need to click on the following link to complete the Temple St. Ratings:**

https://www.surveymonkey.com/s/7XSL8L3

For example, to rate Video 1 you will need to open the video on the memory key (using the attached instructions), then watch the video, click on the above link, and finally choose the appropriate rating for Video 1. When you have completed Video 1, click ‘Next’ on the survey, open Video 2 on the memory key, watch the video and choose the appropriate rating for Video 2. Please continue to complete the ratings in this manner until you have finished rating Video 10. If you wish to change your mind about any previous ratings, please do not hesitate to click on ‘Prev’ and work your way back to the appropriate video. To complete the survey click on ‘Done’ when you have finished Video 10.
When listening to the samples, we recommend that you use Sennheiser 90 headphones (or equivalent). Genelec speakers are also appropriate. Where high-quality audio equipment is not available, the research assistants can arrange a loan of Genelec speakers from the Speech and Language –Therapy Department at Temple St Children’s University Hospital. Please do not hesitate to contact either of us if this is the case.

You will be required to complete this first set of ratings to by the **21/02/2014**.

When you have completed Part 1 of the study, we will then send you another email outlining details of Part 2 of the study in four weeks’ time. This email will contain visual analogue scales. You will then rate the same 10 samples according to the visual analogue scales. The second set of ratings should also take about one hour.

Please see the attached instructions for opening the videos on the encrypted memory keys (there will also be a hard copy of these instructions enclosed in the package you received with your keys).

Yours sincerely,

Sarah Jane Osborne                      Martina Dwyer                      Professor Triona Sweeney
To all participants,

Many thanks for agreeing to participate in our research study on Hypernasality, comparing the use of the Temple Street Scale of Nasality and Nasal Airflow with the use of a Visual Analogue Scale for rater-reliability in both the trained and untrained listener.

Please note you should receive your encrypted keys in the next day or so with this letter of instruction on how to open the keys. Please see previous attachment for the password and below for instructions of use. You will have received only one of the below keys, either the DataTraveler or the SanDisk. Please refer only to the appropriate instructions.

Please ensure not to copy the files to your own computer. Collection of the keys will be organised at the end of the project to ensure all files have been deleted, where this is not possible you may be asked to return via registered post.

INSTRUCTIONS FOR OPENING KEYS:

**Data Traveler (Silver Key)**

1. Open the key
2. Run DTLplus_launcher.exp
3. Data Traveler icon will appear (this may take a minute or two to load)
4. Enter password (as per e-mail).
5. Open folder to view files.

(Please note if your key doesn’t load as above go into my computer; two drives will appear, one is the Launcher (shield icon) and the second contains the files, you need to use the Launcher first to enter the password, then you may need to go back into my computer and open up the files which our stored under a different drive)

**SanDisk (Slim line Shiny Keys)**

1. Open the key.
2. Click on RunSanDiskSecureAccess_Win.exe (this should be the last file on the drop down menu)
3. “Skip” registration.

4. Enter password (as per e-mail)

5. Click OK twice to close irrelevant windows.

6. You should see the files in the remaining open window.

If you have any issues opening the memory keys or any other queries, please do not hesitate to contact us. Thank you for your continued participation in this project.

Yours sincerely,

Sarah Jane Osborne          Martina Dwyer          Professor Triona Sweeney

---

Company Registration N0.351404    Charity No. CHY 229
Directors: Mr. Donal Walsh (Chairperson), Ms. Siobhan Brady, Dr. Liam Claffey, Ms. Suzanne Dempsey, Dr. Michael Drumm, Mr. John Fitzpatrick, Mr. Fionn MacCumhaill, Mr. Derek McGrath, Mr. Frank McManus, Sr. Margherita Rock, Mr. Sean
Appendix 9 (a): Demographics questionnaire

1. How long have you been qualified as an SLT?
   - 0 – 2 years
   - 2 – 5 years
   - Over 5 years
   - Over 10 years

2. How long have you been trained in the Temple St. Scale of Nasality and Nasal Airflow Errors?
   - How long have you been trained in the Temple St. Scale of Nasality and Nasal Airflow Errors? 0 – 2 years
   - 2 – 5 years
   - Over 5 years

3. How long have you been working with children with hypernasality?
   - How long have you been working with children with hypernasality? 0 – 2 years
   - 2 – 5 years
   - Over 5 years
   - Over 10 years

4. How many patients did you have on your caseload last year with hypernasality?
   - Less than 30 patients
   - Less than 50 patients
   - Less than 100 patients
   - Over 100 patients

5. In which of the following settings do you work?
   - Community Care
   - Dublin Cleft Centre

6. Which of the following audio equipment are you using to rate the samples?
   - Which of the following audio equipment are you using to rate the samples? Sennheiser 90 headphones or equivalent
   - Genelec speakers
   - If other audio equipment is used, please indicate the model chosen:
     Survey monkey; Presentation of TSS;
Appendix 10 (b); Results from demographics

Q.1)   SLT for;
        Over 5 years;   3 participants (42.86%)
        Over 10 years; 4 participants (57.14%)

Q.2)   Trained in TSS;
        0-2 years;     2 participants (28.57%)
        2-5 years;     3 participants (42.86%)
        Over 5 years; 2 participants (28.57%)

Q.3)   Worked with resonance;
        0-2years;      2 participants (28.57%)
        2-5years;      1 participants (14.29%)
        Over 5years; 2 participants (28.57%)
        Over 10 years;2 participants (28.57%)

Q.4)   No on caseload;
        Less than 10;  2 participants (28.57%)
        Over 100;      5 participants (71.43%)

Q.5)   Work location;
        Community      1 participants (14.29%)
        Dublin Cleft Centre  5 participants (71.43%)

Q.6)   Speakers used;
        Sennheiser 90 headphones or equivalent  3 participants (42.86%)
        Genelec speakers                 3 participants (42.86%)
        (One participant did not respond to Q.6)
Appendix 11; The Temple Street Scale, scale for hypernasality

**Hypernasality:**

- **Present**
  - a) mild, evident but acceptable.
  - b) mild/moderate, 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.

- **Absent**

<table>
<thead>
<tr>
<th>Consistent</th>
<th>Inconsistent</th>
</tr>
</thead>
</table>

Appendix 12; Adaptation of TSS hypernasality scale to Survey Monkey

**Video 1 Hypernasality Rating:**

- Absent
- mild, evident but acceptable.
- mild/moderate, unacceptable distortion, evident on close vowels.
- moderate, evident on close and open vowels.
- moderate / severe, evident on all vowels and some consonants.
- severe, evident on all vowels and most voiced consonants.
**Appendix 13; The Temple Street Scale (Sweeney 2000)**

![Temple Street Scale Logo]

**TEMPLE STREET SCALE OF NASALITY AND NASAL AIRFLOW ERRORS**

- **NAME**
- **AGE**
- **DOA**
- **CLEFT TYPE**

**TESTER**

<table>
<thead>
<tr>
<th>Audio tape</th>
<th>Video tape</th>
<th>Nasometry</th>
<th>PERCI SARS</th>
</tr>
</thead>
</table>

AV CONSENT OBTAINED:
The puppy is playing with a rope.
The phone fell off the shelf.
Harry is driving a car.
Tim's hand is cleaner than the other.
Tom saw a film in the year.
A tail is like a helical.
Tom is putting on a hat.
Dadie unwound the door.
I put some pepper on a bag.
The eggs are in the box.
Sara is washing a dish dish.
Chele is watching a football match.
John's got a single bedtime.
The young plants are rising.
The bell's ringing.
Karen is making a cake.
George's got a bag of legs.
Mum had her head.
Chester sniffed something crumbled up his floor.
We were over all year.
Laure will wear a yellow shade.
Temple Street Scale - Nasal Airflow

Nasal Emission:
- Frequent
- Infrequent

Nasal Turbulence:
- Frequent
- Infrequent

Nasal Velopharyngeal Intensity:
- Present
- Absent

Mirror Test: R /p t k/ L /p t k/

Nasal Obstruction:
1. Nasal flare
2. Nasal Granuloma
3. Nasal Granuloma

INSTRUMENTAL SCORES

Pressure Flow: NS VPA NS VPA Nasalance: Total HD LP

CLEFT TYPE SPEECH CHARACTERISTICS: (Adapted from GOI SPAS, Dell, Harding & Greenwald, 1977)

JACK & JILL & HUMpty DUMpty
Corrected: 1 - 20, 60 - 70

Conversational speech

Temple Street Scale - Nasal

Hypopharynx:
- Present
- Absent

a) Mild, evident but acceptable
b) Moderate, unacceptable distortion, evident on close vowels
c) Severe, evident on all vowels and some consonants

Consistent

Inconsistent

Cat & Dog
- Present
- Absent

a) Mild - evident but acceptable
b) Moderate - all vowels reduced nasality
c) Severe - total nasalization of nasal consonants
| Consonant Production | m | p | b | f | v | θ | d | s | z | θ | l | b | z | w | l | j |
|----------------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 52                   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| 59                   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

**Dysphagia/Stimulability**

| m | p | b | f | v | θ | d | s | z | θ | l | b | z | w | l | j |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

**Developmental Errors:**

- **Voice**: Normal, Dysphonic, Reduced Volume
- **Respiratory**: Normal, Loss
- **Language**: Normal, Delayed, Disordered

**Craniofacial** (Adapted from GOS S.PASS, Sal, Harding & Conwall 1997):

1. max: NAD deviated septum obstruction aber anos
2. lip: NAD restricted movement open posture
3. occlusion: I II III open bite
4. deviation: NAD superimposition missing teeth malaligned
5. tongue: mobility passive hanging tie
6. palatal form: present absent
7. palatal size: 1 < 2mm 2 3-5mm 4-8mm 5 complete breakdown
8. palatal mobility: normal
9. nasopharynx - tonsils deep pharynx pharyngeal wall movement pharyngoplasty

**Surgical Status:**

**Speech & Language Therapy:**

- 0 Not indicated
- 1 Therapy ongoing
- 2 Regular review
- 3 Group
- 4 Unsuitable
- 5 Short term
- 6 No uptake
- 7 Discharge

**Influencing Factors**

**Recommendations:**

© Triona Sweeney, The Children’s University Hospital, Temple Street, Dublin.
Appendix 14; Copy of VAS instructions (as sent via e-mail)

To all participants,

Thanks again for your continued participation in the project. We really value your commitment to the project so far. We would be extremely grateful if you could complete the final part of the project by the 31st of March.

To complete the final part of the project:

1. Open the same videos as before on your encrypted keys (please do not hesitate to contact us if there are any further issues with any of the keys).
2. Print off a copy of the attached visual analogue scales.
3. For each video, make a mark on the line anywhere from 'Normal' to 'severely hypernasal' to make a visual analogue rating (using a pen or pencil)
4. To return the scales you can either scan your completed ratings or attach the scanned file in an email to us at this address, or if you prefer, you can post your ratings to the following address:

Sarah Jane Osborne c/o Jacinta Shanahan
University of Limerick Speech and Language Therapy Clinic,
Clinical Therapies Dept.,
Health Science Building,
University of Limerick,
Castletroy,
Limerick

If you have any queries on how to make the ratings or any further technical problems, please let us know.

Kindest regards,

Sarah Jane Osborne and Martina Dwyer.

Company Registration No.351404    Charity No. CHY 229
Directors: Mr. Donal Walsh (Chairperson), Ms. Siobhan Brady, Dr. Liam Clafley, Ms. Suzanne Dempsey, Dr. Michael Drumm, Mr. John Fitzpatrick, Mr. Fionn MacCumhaill, Mr. Derek McGrath, Mr. Frank McManus, Sr. Margherita Rock, Mr. Sean
Appendix 15; Visual Analogue Scale

Video 1:  
[[-----------------------------]]  
Normal  Severe hypernasal

Video 2:  
[[----------------------------]]  
Normal  Severe hypernasal

Video 3:  
[[----------------------------]]  
Normal  Severe hypernasal

Video 4:  
[[----------------------------]]  
Normal  Severe hypernasal

Video 5:  
[[----------------------------]]  
Normal  Severe hypernasal

Video 6:  
[[----------------------------]]  
Normal  Severe hypernasal

Video 7:  
[[----------------------------]]  
Normal  Severe hypernasal

Video 8:  
[[----------------------------]]  
Normal  Severe hypernasal

Video 9:  
[[----------------------------]]  
Normal  Severe hypernasal