Comparison of nasalance scores obtained with the Nasometers 6200 and 6450

Gillian de Boer MA MSc and Tim Bressmann PhD

Gillian de Boer is a master's graduate in the Department of Speech-Language Pathology, University of Toronto, and Tim Bressmann is an associate professor in the Department of Speech-Language Pathology, University of Toronto, Toronto, Ontario.

Address correspondence to:

Tim Bressmann, Ph.D.

Department of Speech-Language Pathology

University of Toronto

160-500 University Avenue

Toronto, ON M5G 1V7

Canada

Tel. ++1-416-978-7088, Fax -1596

tim.bressmann@utoronto.ca

This research was supported by an Operating Grant from the Canadian Institutes of Health Research (grant fund number 485680).

Running title: Comparison of Nasometers 6200 and 6450

Comparison of nasalance scores obtained with the Nasometers 6200 and 6450

Abstract

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

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

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

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

KEYWORDS: acoustics, nasality, resonance, speech production, nasometry

Introduction

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

The most commonly used instrument for the measurement of nasalance is the Nasometer by KayPENTAX (KayPENTAX, Montvale, NJ). Different generations of the Nasometer have used different methods of signal processing. The original Nasometer 6200, first introduced in 1986, converted the sound pressure levels to direct current and sent calculations to the computer. It was not possible to record an audiofile with the Nasometer 6200. A special innovation of the Nasometer 6200 was that it used a band pass filter with a centre frequency of 500 Hz and a range of 300 Hz to emphasize the extra resonances in this frequency range that are associated with hypernasal speech (Fletcher and Bishop, 1970). Baken and Orlikoff (2000) have argued that this filter range was chosen *a priori* without a sound rationale. Awan (1998) has pointed out that the signal filtering affects the researchers' ability to do acoustic analyses on nasalance recordings. Notwithstanding these criticisms, subsequent Nasometer models have employed the same filter

characteristics in the signal processing. The second generation Nasometer II 6400, first distributed in 2002, used a preamplifier together with a dedicated computer soundcard. The latest model, the Nasometer II 6450, was introduced in 2009. It uses an external universal serial bus sound card and transfers a digital sound file to the computer.

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

Another goal of the current study was to assess the test-retest variability for the new Nasometer 6450 in comparison to the Nasometer 6200. The initial studies of immediate test-retest variability in nasalance scores of normal participants, measured with the Nasometer 6200, found that between 94-100% of recordings were within 3 nasalance points for three repeated recordings (Seaver et al., 1991; Litzaw and Dalston 1992). Using oral and balanced sentences, variability among hypernasal participants was found to be greater (Watterson et al., 2006). The immediate test-retest nasalance score variability of the Nasometer 6200 was within 5 points for 88% of the participants. When the headset and microphone were removed and replaced between readings,

89% of the participants were within 9 points. The authors suggest that, in hypernasal participants, a test-retest variability of 10 nasalance points may be the normal range of variability. However, the authors point out that, even when the headset was removed and replaced, 61% of the participants were still within 5 nasalance points (Watterson et al., 2006).

In a study of nasalance testing materials for normal female speakers of Cantonese, Whitehill (2001) measured day-to-day variability. For oral stimuli (a sentence and paragraph), 95% of the returning participants were within 5 points of their previous score. For balanced stimuli, 93% of participants were within 8 points and for nasal stimuli, 96% of participants were within 10 points. Lewis et al. (2008) compared immediate and long-term test-retest variability for normal participants using the Nasometer 6400. Long-term variability was greater but did not increase with the time between measurements. Re-testing was done on the same day, twice a day for five consecutive days, and once a week for three weeks. The stimuli consisted of the Turtle (oral) and the Mouse (balanced) passages. The short-term variability was less than 5 nasalance points for over 90% of the participants. However, the long-term variability of 90 % of the participants was within a range of 6 to 8 nasalance points.

These studies provide substantial information about the variability of nasalance scores obtained with the Nasometers 6200 and 6400. However, with the introduction of the new Nasometer 6450, it becomes important to assess the test-retest characteristics for this instrument and to compare it to the original Nasometer 6200. Also, knowledge of normal variation of nasalance scores obtained from nasal stimuli over time is limited to the study by Whitehill (2001) who demonstrated that nasal test stimuli had higher test-retest variability. In a similar vein, Watterson

et al. (2006) found that hypernasal speakers had higher test-retest variability. This suggests that there could be a relationship between subjects' nasalance scores and their mean test-retest variability.

Zajac et al. (1996) investigated the relationship between the sensitivity of the microphones used in the Nasometer 6200 and found that the equipment characteristics influenced the nasalance scores. Other than this study, almost all of the research concerning nasometer measurements has been carried out with human participants. Synthetic test sounds have the advantage that they do not have the variability of human speakers. We therefore argue that it would also be desirable to characterize the new Nasometer 6450 in more detail using synthetic test sounds.

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

Study 1: Comparative assessment of the Nasometers 6200 and 6450 with square wave test sounds

Methods

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

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

The Nasometer was calibrated using a single loudspeaker or sound source. The previous research by Zajac et al. (1996) also used a single loudspeaker to assess different Nasometer headsets. In the present study, a stereo configuration with two loudspeakers addressing the two microphones on the Nasometer headset was used. This was done to estimate the crossover between the two microphones and to assess the robustness of the nasometer measurements. Audio editing software was used to create five different stereo panoramas for each sound file: hard left, three-quarter left, balanced, three-quarter right, and hard right. In theory, this should lead to measurements approximating 0%, 25%, 50%, 75% and 100% nasalance respectively. Due to

acoustic spillover around the 25 dB separation plate, the real measures cannot be expected to be so clear cut (KayPENTAX, 2010). Nevertheless, it was expected that the nasalance scores for the hard left and three-quarter left recordings were exact mirror images of the hard right and three-quarter right recordings.

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

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

The loudspeakers and the Nasometer headset were placed in a custom-made holder (Figure 1). The holder was made from blocks of packing foam. There was a 1 cm foam spacer between the loudspeaker and the Nasometer separator plate on either side. The loudspeaker cones were 2 cm

beneath the Nasometer microphones. In the default setup, the nasal microphone recorded from the right loudspeaker and the oral microphone from the left loudspeaker. To control for differences in the sound energy emitted from the loudspeakers themselves, the orientation of the setup was switched for a part of the recordings so that the nasal microphone recorded from the left loudspeaker and the oral microphone from the right loudspeaker.

[Figure 1 about here]

Three repeated nasalance recordings were made for each frequency in the hard left, three-quarter left, balanced, three-quarter right, and hard right stereo panoramas. With the Nasometer separation plate and the microphones reversed, the test sounds in the three-quarter right, three-quarter left and balanced panoramas were repeated. Nasalance scores were recorded three times for each frequency, sound panorama and Nasometer model, resulting in a total of 912 nasalance scores. The Nasometer 6200 provides nasalance scores to two decimal places while the Nasometer 6450 provides integers. Before mean values and standard deviations were calculated, the nasalance scores of the Nasometer 6200 were rounded to the nearest integer using the rounding function in SPSS. The same operation was performed by Watterson et al. (2005) when they compared the Nasometers 6200 and 6400.

Both Nasometers were used with their respective headsets. Before the recordings, the Nasometers 6200 and 6450 were calibrated according to the manufacturer's instructions. The calibration readings were 49.65% for the Nasometer 6200 and 0.96 for the Nasometer 6450. Both calibration scores were within the manufacturer's recommended range.

Results

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

[Figure 2 about here]

The mean nasalance values for the results for all frequencies tested are displayed in Tables 1 and 2. A first inspection of the results drew our attention to the mean nasalance scores that were obtained for the three-quarter right, three-quarter left and balanced sound files for the 755 Hz square wave, measured with the Nasometer 6200. Statistical analysis confirmed that these scores were outliers. The same sound file did not present different results for the Nasometer 6450.

[Tables 1 and 2 about here]

A repeated-measures ANOVA was run for loudspeaker panorama (eight loudspeaker stereo panoramas), nasometer model (6200 vs. 6450) and repetition (three repetitions). Where sphericity was violated, the Greenhouse-Geisser correction was used. There were significant effects for panorama ($p < .001$, $F(1.465, 126) = 3192.752$), nasometer ($p < .005$, $F(1, 18) = 18.419$) and the panorama-nasometer interaction ($p < .01$, $F(2.159, 38.857) = 5.836$). The effect of repetition and the remaining interactions were not significant.

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

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

Study 2: Comparative assessment of the Nasometers 6200 and 6450 with normal participants

Participants

Twenty-five females were recruited from the student population of the University of Toronto's Speech-Language Pathology program. The participants were between 22 and 30 years of age ($M = 24.16$, $SD = 2.375$) and spoke English with the accent that is common to Southern Ontario. All participants provided informed consent. Each participant completed six recording sessions between May and August 2011. Twenty-one of the participants completed their sessions within 7 weeks from start of the study. Due to scheduling conflicts, four participants' recordings took up to 11 weeks to complete. The average time between the first and the final recording was 35 days (range 16 to 77 days).

Methods

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

We used oral, balanced (oral-nasal) and nasal stimuli for the recordings of the participants on the two Nasometers. The stimuli were abbreviated versions of the Zoo Passage (Fletcher, 1976), the Rainbow Passage (Fairbanks, 1963) and the Nasal Sentences (Fletcher, 1976). We used the first two sentences of the Zoo Passage ("Look at this book with us. It's a story about a zoo."), the second sentence of the Rainbow Passage ("The rainbow is a division of white light into many beautiful colors.") and the first sentence from the Nasal Sentences ("Mama made some lemon jam."). In previous research, Bressmann (2005) compared these abbreviated versions of the stimuli to the full versions. The abbreviated versions were used to obtain as many measurements as possible in the limited time the student participants were available. Approval for this study was granted by the University of Toronto's Office of Research Ethics.

Statistical analysis

The nasalance scores of the Nasometers were compared with a repeated measures ANOVA and post-hoc t-tests. The test-retest variability between measurements was examined by calculating the absolute differences in scores between sessions and creating cumulative frequency tables. As

the absolute differences in nasalance scores were not normally distributed, the Wilcoxon signed-rank test was used to determine if some stimuli had greater variance than others. To assess the relationship between the subjects' nasalance scores and their mean test-retest variability, we calculated Pearson product-moment correlation coefficients.

Results

The mean nasalance scores were calculated by averaging the scores for each stimulus for each Nasometer over the six sessions. The results are displayed in Table 3.

[Table 3 about here]

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

Paired t-tests were used to compare the two nasometers post hoc. The mean nasalance score for the Nasometer 6450 (36.47, SD 21.686) was significantly higher ($p < .05$) than the mean of the Nasometer 6200 (35.56, SD 20.885). The mean nasalance scores of the oral, balanced and nasal

stimuli (13.16, SD 6.121; 32.22, SD 5.800; and 62.67, SD 6.267, respectively) differed significantly (all differences $p < .01$). For the sentence-nasometer interaction effect, paired t-tests showed that the Nasometer 6450 scored significantly higher than the Nasometer 6200 for the balanced stimulus ($p < .05$) and the nasal stimulus ($p < .01$). For the effect of session, the paired t-tests indicated that the mean nasalance score of the sixth session (37.10, SD 5.090) differed significantly from the first (35.43, SD 5.308), second (35.560, SD 5.364) and third sessions (35.66, SD 4.305) (all differences $p < .05$).

The variability of nasalance scores over time was assessed by calculating the differences in scores between the six sessions for each sentence, nasometer and participant. The differences were converted to absolute values, and the mean and standard deviation were calculated. These results are displayed in Table 4. Wilcoxon signed-rank tests were conducted to compare the mean differences for the different stimuli. There were significant differences between the nasal stimulus and the oral ($z = -2.678$, $p < .01$) and balanced stimuli ($z = -2.886$, $p < .01$) for the Nasometer 6200. For the Nasometer 6450, there was a significant difference between the nasal and oral stimuli ($z = -3.116$, $p < .01$). There were no significant differences between the machines.

[Table 4 about here]

The cumulative frequencies of the absolute differences in nasalance points across six sessions are also shown in Table 4. For both Nasometers, 90.7% of the participants' nasalance scores for the oral stimulus were within 6 points. For the balanced sentence, 92.3% (Nasometer 6200) and

93.3% (Nasometer 6450) of the scores were within 7 points of the participants' other scores. An 8 point spread captured 91.5% (Nasometer 6200) and 91.7% (Nasometer 6450) of nasalance scores for the nasal sentence.

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

DISCUSSION

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

The initial analysis of the sound pressure levels transmitted by the two loudspeakers for the synthetic test sounds demonstrated that the left loudspeaker was slightly louder. These differences were probably due to minor differences in the manufacturing of the two loudspeakers. The mean nasalance scores obtained from the hard left and hard right sound files, 7 and 93% respectively, suggest that 7% of the acoustic energy presented on one side of the sound separator plate makes its way to the opposite microphone. According to the manufacturer, the sound separator plate attenuates sound transmission to the opposite microphone by 25 dB (KayPENTAX, 2010). Therefore, the nasalance scores will never be zero or 100 (Gildersleeve-Neumann and Dalston, 2001). It also follows that the mean values for the loudspeaker balances

can only roughly approximate the theoretically expected scores of 0%, 25%, 50%, 75% and 100% nasalance for the hard left, three-quarter left, balanced, three-quarter right, and hard right stereo panoramas, respectively.

With the test frequency of 755 Hz, our Nasometer 6200 produced nasalance scores that deviated from the scores for the other frequencies by approximately 11 to 19 points. This effect was reproducible and persisted when the microphones were reversed. Since the same effect was not observed for the Nasometer 6450, it did not appear that this particular frequency excited resonances in any elements of the recording contraption. It would be interesting to study whether this effect was specific to our particular Nasometer 6200 or if other instruments show a similar phenomenon for this frequency. Zajac et al. (1996) demonstrated differences in the measurements obtained with different headsets for the Nasometer 6200. Both Nasometers were used with the headsets that they were delivered with, so the possibility of a mismatch in microphone sensitivity cannot be ruled out.

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

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

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

The mean differences in scores between sessions for the oral and balanced stimuli were comparable to those found by Lewis et al. (2008) using the Nasometer 6400 while the mean

differences for the nasal materials were comparable to Whitehill (2001). The trend of increasing variability of scores with increased nasal content that was evident in the data reported by Whitehill (2001) was confirmed in the present study. This increase may be a simple effect of the averaging of the nasalance score because a stimulus containing nasal sounds will have nasalance spikes in the nasalance trace. Such statistical outliers can skew the frequency distribution and distort the arithmetic mean (Ferguson, 1981). While there were significant differences between the variability of the stimuli, there were none between the nasometers.

For clinical practice, the distribution of the differences in nasalance points is more meaningful than the mean differences and standard deviations. The cumulative frequency tables illustrated an increase in measurement variability as the nasal content of the test stimuli increased. For both the Nasometer 6200 and the Nasometer 6450, a 6 point spread captured 90.7% of the differences in scores for the oral stimuli. A 7 point spread captured 92.3% of the differences for the balanced stimulus for the Nasometer 6200 and 93.3% for the Nasometer 6450. These results are very similar to what Lewis et al. (2008) found for these stimuli with the Nasometer 6400. For the nasal sentence, an 8 point spread accounted for 91.5% of the differences for the Nasometer 6200 and for 91.7% for the Nasometer 6450. This was comparable to the results by Whitehill (2001) who found that 89.3% of her participants were within 8 points. Given these results, clinicians would be advised to allow as much as 6 nasalance points for an oral stimulus and 8 nasalance points for a nasal stimulus when using a nasometer to assess possible changes in a patient's oral-nasal balance.

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

Conclusion

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

Acknowledgement

The authors wish to thank Dr. Pascal van Lieshout and Dr. Kiran Kulkarni for their help and guidance with this research project. We also thank two anonymous reviewers and the Associate Editor of the Cleft Palate-Craniofacial Journal for their constructive and thorough reviews. Parts of this paper were presented as a poster at the 69th annual meeting of the American Cleft Palate Association, April 2012 in San Jose, CA.

References

- Awan SN. Analysis of nasalance: NasalView (the Nasalance Acquisition System). In: Ziegler W, Deger K, eds. *Clinical phonetics and linguistics*. London: Whurr; 1998: 518-527.
- Awan SN, Omlar K, Watts CR. Effects of computer system and vowel loading on measures of nasalance. *J Speech Lang Hear Res*. 2011; 54: 1284-1294.
- Baken RJ, Orlikoff RF. *Clinical measurement of speech and voice, 2nd edition*. Clifton Park, NY: Thomson; 2000.
- Bressmann T. Comparison of nasalance scores obtained with the Nasometer, the NasalView, and the OroNasal System. *Cleft Palate Craniofac J*. 2005; 42: 423-433.
- Dalston RM, Warren DW, Dalston ET. Use of nasometry as a diagnostic tool for identifying patients with velopharyngeal impairment. *Cleft Palate Craniofac J*. 1991a; 28: 184-188.
- Dalston RM, Warren DW, Dalston ET. A preliminary investigation concerning the use of nasaometry in identifying patients with hyponasality and/ or nasal airway obstruction. *J Speech Hear Res*. 1991b; 34: 11-18.
- Fairbanks G. *Voice and Articulation Drillbook, 2nd edition*. New York, NY: Harper and Bros.; 1960.

- Ferguson GA. *Statistical analysis in psychology and education, fifth edition*. New York, NY: McGraw-Hill; 1981.
- Fletcher SG, Bishop ME. Measurement of nasality with TONAR. *Cleft Palate J*. 1970; 7: 610-621.
- Fletcher SG. Nasalance vs. listener judgements of nasality. *Cleft Palate J*. 1976; 13: 31-44.
- Gauster A, Yunusova Y, Zajac D. The effect of speaking rate on velopharyngeal function in healthy speakers. *Clinical Linguist Phon*. 2010; 24: 576-588.
- Gildersleeve-Neumann CE, Dalston RM. Nasalance Scores in Noncleft Individuals: Why Not Zero? *Cleft Palate Craniofac J*. 2001; 38: 106-111.
- Karnell MP. Nasometric discrimination of hypernasality and turbulent nasal airflow. *Cleft Palate Craniofac J*. 1995; 32: 145-148.
- KayPENTAX. Instruction manual Nasometer II Model 6450, issue C. Lincoln Park, NJ: KayPentax; 2010.
- Kummer A. Nasometry. In: Kummer A, ed: *Cleft Palate and Craniofacial Anomalies – Effects on Speech and Resonance, 2nd edition*. New York, NY: Delmar Cengage Learning; 2008: 376-414.

- Lewis KE, Watterson T, Blanton A. Comparison of short-term and long-term variability in nasalance scores. *Cleft Palate Craniofac J*. 2008; 45: 495-500.
- Litzaw LL, Dalston RM. The effect of gender upon nasalance scores among normal adult speakers. *J Commun Disord*. 1992; 25: 55-64.
- Olsson P, Bende M. Influence of environmental temperature on human nasal mucosa. *Ann Otol Rhinol Laryngol*. 1985; 94: 153-155.
- Seaver EJ, Dalston RM, Leeper HA, Adams LE. A study of nasometric values for normal nasal resonance. *J Speech Hear Res*. 1991; 34: 715-722.
- Watterson T, Lewis K, Brancamp T. Comparison of nasalance scores obtained with the Nasometer 6200 and the Nasometer II 6400. *Cleft Palate Craniofac J*. 2005; 42: 574-579.
- Watterson T, Lewis KE. Test-retest nasalance score variability in hypernasal speakers. *Cleft Palate Craniofac J*. 2006; 43: 415-419.
- Whitehill TL. Nasalance measures in Cantonese-speaking women. *Cleft Palate Craniofac J*. 2001; 38: 119-125.

Zajac DJ, Lutz R, Mayo R. Microphone sensitivity as a source of variation in nasalance scores. *J Speech Lang Hear Res.* 1996; 39: 1228-1231.

Figure legends

Figure 1. Measuring set-up for square wave test sounds.

Figure 2. Sound pressure level in dB (C) by frequency for the left and right loudspeaker.

Table legends

Table 1. Nasalance scores for square wave test sounds in different stereo panoramas, recorded with the Nasometer 6200.

Table 2. Nasalance scores for square wave test sounds in different stereo panoramas, recorded with the Nasometer 6450.

Table 3. Nasalance scores by Nasometer model and session.

Table 4. Means and Cumulative Frequencies of Test-Retest Differences by Nasometer Model.

Table 1. Nasalance scores for square wave test sounds in different stereo panoramas, recorded with the Nasometer 6200.

Frequency	Left		$\frac{3}{4}$ Left		Balanced		$\frac{3}{4}$ Right		Right		Rev $\frac{3}{4}$ Left		Rev Balanced		Rev $\frac{3}{4}$ Right	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
805	11	0	25	0	47	0	72.33	0.58	88	0	76	0	54	0	29	0
755	7.33	0.58	42	0	47	0	60.67	0.58	91.67	2.31	60	0	54.67	0.58	41	0
705	9	0	23	0	47	0	72	0	90	0	79	0	56	0	30.67	0.58
655	9	0	24	0	46	0	71	0	90	0	78	0	55.33	0.58	31	0
605	7.33	0.58	25	0	46.33	1.15	71	0	89	0	78	0	55	0	31	0
555	10	0	22.33	0.58	47	0	74	0	89	0	80.67	0.58	56	1.73	29.33	0.58
505	8	0	18	0	45	0	74	0	92	0	83.67	0.58	58.33	0.58	28.67	0.58
455	5	0	22	0	47.33	0.58	71.33	0.58	94.67	0.58	77.33	0.58	55	0	31	0
405	7	0	24.33	0.58	47	0	71	0	94.67	0.58	77.67	0.58	54.33	0.58	30.33	1.15
355	4	0	22	0	47.33	0.58	73	0	94.67	0.58	78.67	0.58	54	0	29.67	0.58
305	5	0	22	0	45.67	0.58	74	0	93	0	79	0	51.67	0.58	26.67	0.58
280	7	0	17.67	0.58	45.67	0.58	75	0	92	0	81	0	53.33	0.58	26	0
255	4	0	20.67	0.58	46	0	72.33	0.58	96	0	76.67	0.58	54	0	28.67	0.58
230	3	0	22	0	45	0	71.33	0.58	95.33	0.58	78	0	54	0	27	0
205	4	0	22	0	46	0	74	0	95	0	81	0	55	0	28	0
180	5	0	21.67	0.58	48	0	75	0	95	0	82	0	57	0	29	0
155	7	0	21	0	48	0	76	0	95	0	84	0	58	0	28	0
130	8	0	24	0	47.67	0.58	74	0	94	0	82	0	58.67	0.58	30.67	0.58
105	6	0	22	0	46	0	72	0	94.33	0.58	79	0	55.33	1.15	29	0
Grand Mean/ sd	6.67	2.23	23.19	4.96	46.58	0.92	72.32	3.21	92.81	2.53	78.51	5.02	55.25	1.78	29.72	3.12

Table 2. Nasalance scores for square wave test sounds in different stereo panoramas, recorded with the Nasometer 6450.

Frequency	Left		$\frac{3}{4}$ Left		Balanced		$\frac{3}{4}$ Right		Right		Rev $\frac{3}{4}$ Left		Rev Balanced		Rev $\frac{3}{4}$ Right	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
805	9	0	24	0	46	0	70.33	0.58	94	0	77	0	52	0	27	0
755	7.33	0.58	26	0	47	0	70.33	0.58	93.67	0.58	72.33	0.58	51	0	28	0
705	6.33	0.58	23	0	47	0	72	0	93	0	76	0	52	0	27	0
655	6	0	22	0	44.67	0.58	71	0	95	0	77	0	53	0	28	0
605	9	0	21	0	45	0	72	0	93	0	79	0	53	0	26.67	0.58
555	7.33	0.58	20	0	46	0	74	0	92	0	80	0	52	0	23	0
505	6	0	17.33	0.58	42	0	69	0	96.67	0.58	77	0	54	0	27	0
455	4	0	23	0	46	0	71	0	95.33	0.58	79.33	0.58	55	0	29	0
405	10	0.58	22	0	46	0	73	0	91.33	0.58	76	0	53	0	28	0
355	5	0	18	0	45	0	75	0	92	0	81	0	53	0	23	0
305	7	0	18	0	42	0	70	0	95	0	76	0	50.67	0.58	24.67	0.58
280	7	0	20	0	46	0	73	0	96	0	80	0	53	0	24.33	0.58
255	8	0	19.33	0.58	46	0	74	0	93	0	80	0	54	0	26	0
230	7	0	21	0	47	0	75	0	92	0	81	0	54	0	25	0
205	8	0	20	0	46	0	74	0	93	0	82	0	56	0	27	0
180	8	0	17	0	45	0	74	0	94.33	0.58	83	0	57	0	28	0
155	7.33	0.58	18	0	46	0	75	0	93	0	81	0	55	0	26	0
130	7.33	0.58	17	0	45	0	74	0	91.33	0.58	81	0	55	0	26	0
105	8.33	0.58	19	0	44	0	71	0	94.33	0.58	78	0	53	0	27	0
Grand Mean/ sd	7.26	1.42	20.30	2.55	45.35	1.43	72.51	1.91	93.58	1.55	78.77	2.66	53.46	1.64	26.35	1.70

Table 3. Mean Nasalance Scores by Nasometer and session.

Session	Nasometer 6200						Nasometer 6450					
	Oral		Balanced		Nasal		Oral		Balanced		Nasal	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	12.00	4.967	31.24	5.418	61.60	6.298	12.84	7.087	31.92	6.745	62.96	6.871
2	12.72	4.641	31.28	5.144	61.28	6.655	12.28	7.009	32.88	6.882	62.92	6.788
3	12.60	5.220	31.12	4.702	61.32	5.258	12.60	5.354	32.80	5.657	63.52	5.810
4	13.12	6.772	31.48	5.782	61.76	6.372	12.84	6.183	32.36	5.816	63.84	5.669
5	14.12	7.161	32.08	6.770	62.52	6.603	13.76	6.346	32.72	5.792	63.56	6.443
6	14.68	6.369	32.88	5.102	62.36	6.217	14.40	6.357	33.84	5.963	64.44	6.449
Grand Mean / SD	13.21	5.904	31.68	5.468	61.81	6.164	13.12	6.349	32.75	6.085	63.54	6.269

Table 4. Means and Cumulative Frequencies of Test-Retest Differences by Nasometer Model.

Nasalance Score Difference	Nasometer 6200			Nasometer 6450		
	Oral	Balanced	Nasal	Oral	Balanced	Nasal
	frequency / %	frequency / %	frequency / %	frequency / %	frequency / %	frequency / %
0-5	315 / 84.0	309 / 82.4	271 / 72.3	314 / 83.7	308 / 82.1	276 / 73.6
≤ 6	340 / 90.7	332 / 88.5	304 / 81.1	340 / 90.7	332 / 88.5	304 / 81.1
≤ 7	347 / 92.5	346 / 92.3	329 / 87.7	355 / 94.7	350 / 93.3	324 / 86.4
≤ 8	354 / 94.4	359 / 95.7	343 / 91.5	364 / 97.1	364 / 97.1	344 / 91.7
≤ 9	357 / 95.2	370 / 98.7	352 / 93.9	372 / 99.2	369 / 98.4	354 / 94.4
≤ 14	375 / 100	375 / 100	370 / 98.7	375 / 100	375 / 100	375 / 100
≤ 19			375 / 100			
Mean Difference	2.89	2.99	4.07	2.83	3.20	3.92
SD	1.51	1.163	1.378	1.191	1.122	1.109

Oral microphone

Nasal microphone

Extended
bass
cabinet

Loudspeaker
membrane

Nasometer
sound
separation
plate