JSLHR

Research Note

Objective Measure of Nasal Air Emission Using Nasal Accelerometry

Meredith J. Cler,[a,b] Yu-An S. Lien,[c] Maia N. Braden,[d] Talia Mittelman,[b,c] Kerri Downing,[b] and Cara E. Stepp[a,b,c,e]

Purpose: This article describes the development and initial validation of an objective measure of nasal air emission (NAE) using nasal accelerometry.
Method: Nasal acceleration and nasal airflow signals were simultaneously recorded while an expert speech language pathologist modeled NAEs at a variety of severity levels. In addition, microphone and nasal accelerometer signals were collected during the production of /pɑpɑpɑpɑ/ speech utterances by 25 children with and without cleft palate. Fourteen inexperienced raters listened to the microphone signals from the pediatric speakers and rated the samples for the severity of NAE using direct magnitude estimation. Mean listener ratings were compared to a novel quantitative measurement of NAE derived from the nasal acceleration signals.
Results: Correlation between the nasal acceleration energy measure and the measured nasal airflow was high (r = .87). Correlation between the measure and auditory-perceptual ratings was moderate (r = .49).
Conclusion: The measure presented here is quantitative and noninvasive, and the required hardware is inexpensive ($150). Future studies will include speakers with a wider range of NAE severity and etiology, including cleft palate, hearing impairment, or dysarthria. Further development will also involve validation of the measure against airflow measures across subjects.

Velopharyngeal dysfunction (VPD) occurs as a result of inappropriate opening or closing of the velopharyngeal (VP) port. VPD can be due to structural, neurological, or functional impairments; in particular, 20%–30% of children with cleft palate have VPD even after initial cleft palate repair (Rintala & Haapanen, 1995; Witt, Wahlen, Marsh, Grames, & Pilgram, 1998). VPD has a variety of consequences, including hypernasality, hyponasality, and nasal air emission (NAE; Kummer, 2008). The percept of nasality is well correlated with acoustic measures such as nasalance (e.g., Brancamp, Lewis, & Watterson, 2010; Brunnegård, Lohmander, & van Doorn, 2012). However, because total nasalance scores increase when speech samples are known to include NAE (Karnell, 1995), the presence of NAE reduces the accuracy of nasalance as a correlate of nasality (Dalston, Warren, & Dalston, 1991). This suggests that nasometry cannot differentiate between NAE and hypernasality. There is currently no acoustic measure of NAE used clinically.

NAE, also referred to as *nasal escape* or *nasal emission*, occurs when air leaks from the oral cavity into the nasal cavity, either due to a gap in the VP port at times when the port should be tightly closed or due to an oronasal fistula (Kummer, 2008; Wyatt et al., 1996). NAE is particularly noticeable during voiceless obstruents such as /p t k s f ʃ θ/, in which the speaker attempts to build up pressure in the oral cavity that is needed for intelligible production of the phoneme. As a result, NAE is highly correlated with weak pressure consonants (Dotevall, Lohmander-Agerskov, Ejnell, & Bake, 2002), leading to muffled and indistinct speech (Kummer, 2008), regardless of the perceptual salience of the NAE.

In fact, the perception of NAE is quite varied. NAE can be inaudible, perceptible but quiet, or obvious and distracting, as in the case of nasal turbulence or rustle (Kummer, 2008; Kummer, Curtis, Wiggs, Lee, & Strife, 1992). The auditory perception of NAE severity is affected by a variety of factors, including the total volume of air passing through the VP opening, the degree of airway resistance, and the

[a]Graduate Program for Neuroscience–Computational Neuroscience, Boston University, MA
[b]Department of Speech, Language, and Hearing Sciences, Boston University, MA
[c]Department of Biomedical Engineering, Boston University, MA
[d]Department of Surgery, Division of Otolaryngology, Voice and Swallow Clinics, University of Wisconsin–Madison
[e]Department of Otolaryngology–Head and Neck Surgery, Boston University School of Medicine, MA
Correspondence to Cara E. Stepp: cstepp@bu.edu

Editor and Associate Editor: Kate Bunton
Received November 24, 2015
Revision received January 29, 2016
Accepted February 1, 2016
DOI: 10.1044/2016_JSLHR-S-15-0407

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

size of the VP gap. Thus the relationship between the severity of the structural/functional impairment and the severity of the perceived speech abnormality is nonlinear—making NAE difficult to assess (Kummer, Briggs, & Lee, 2003). In clinical settings, NAE is typically evaluated by expert auditory perception, with or without additional visual inspection of condensation on a mirror under the nose, and rarely, via direct measurement of nasal airflow (Kummer, Clark, Redle, Thomsen, & Billmire, 2012). Clinician reliability in perceptually detecting these subtle sounds is heavily influenced by training and experience (Brunnegård, Lohmander, & van Doorn, 2009), despite efforts to construct meaningful perceptual scales (Baylis, Chapman, Whitehill, & The Americleft Speech Group, 2015; Baylis, Munson, & Moller, 2011; Brunnegård et al., 2009; John, Sell, Sweeney, Harding-Bell, & Williams, 2006; McWilliams & Philips, 1979). Accurate assessment of NAE is vital for treatment planning and progress tracking in this population. Thus, here we present the development and initial validation of a proposed objective measure of NAE.

If air is released through the VP port, either due to typical nasalization, hypernasality, or NAE, a nasal accelerometer can detect the mechanical response of the tissues to the nasal airflow. Although nasal accelerometry has been used in other measures of VP function (Mra, Sussman, & Fenwick, 1998) and some measures derived from nasal acceleration correlate well with perceptual ratings of hypernasality (Laczi, Sussman, Stathopoulos, & Huber, 2005; Redenbaugh & Reich, 1985), it has not previously been used to detect NAE. To be specific, the proposed measure is based on the energy detected by a nasal accelerometer; thus we will refer to it as the nasal acceleration energy measure, or NAEM. The NAEM uses nasal accelerometer measurements during the hold period of stop consonants (here, /p/), allowing it to isolate NAE from typical nasalization or hypernasality, which are often measured during vowels or over running speech. In addition, the NAEM provides a measure of NAE severity. This differentiates it from, for example, the N-RamP measure (Bunton, Hoit, & Gallagher, 2011), which is used to detect whether a particular user has NAE or not without a severity rating. This article will compare the NAEM to (a) directly measured nasal airflow in an expert speech language pathologist (SLP) mimicking different levels of NAE and (b) listener perception of NAE in a cohort of pediatric speakers with and without a history of cleft palate.

Method

Speech Samples

Speech samples were provided by an expert adult SLP participant and 25 pediatric participants with and without cleft palate (mean age: 8.4 years; eight girls). Seventeen pediatric speakers were recruited from children with a history of cleft palate undergoing a speech-pathology evaluation at University of Wisconsin–Madison Pediatric Cleft Lip and Palate/Craniofacial Anomalies Clinic at American Family Children's Hospital. Participants (four girls) ranged from 4 to 16 years of age (M = 7.9 years), and had a history of cleft palate (five participants), cleft lip and palate (nine participants), cleft lip and palate with fistula (two participants), or submucosal cleft palate (one participant). Eight children (four girls) with typical development were recorded by researchers at Boston University, and ranged from 4 to 14 years of age (M = 9.5 years). In compliance with the University of Wisconsin Health Sciences Institutional Review Boards (IRBs) and the Boston University IRB, informed consent was obtained from a parent of each child and from the expert SLP. In addition, informed consent, written assent, or verbal assent was obtained from each participant as age appropriate and in compliance with the relevant IRB. Verbal dissent was respected from all participants.

Participants were asked to produce utterances from the MacKay–Kummer Simplified Nasometric Assessment Procedures Test (SNAP test), including connected strings of /pɑ/ tokens (e.g., /pɑpɑpɑpɑpɑpɑpɑ/; MacKay and Kummer, 1994). The expert SLP produced three strings of seven /pɑ/ tokens at each of three specific levels of self-perceived NAE: mild NAE, moderate NAE, and severe NAE. Pediatric participants were instructed to speak normally. Although instructed to produce approximately seven /pɑ/ tokens, utterances ranged from three to 10 tokens.

Speech samples at both sites were recorded using identical hardware and software. Samples were recorded using Audacity software, an external soundcard (Native Instruments Komplete Audio 6 interface; Native Instruments GmbH, Berlin, Germany) and custom instrumentation to measure nasal acceleration and total acoustic output. Nasal acceleration was measured using a wide-band BU Series 21771 accelerometer (Knowles Electronics, Itasca, IL) attached to the participant's nose with double-sided tape. Total acoustic output was measured with a standard headset microphone (Sennheiser PC131) placed approximately 6–10 cm from the mouth at a 45° angle from the midline. All signals were sampled at 44100 Hz. In addition, nasal airflow using a Phonatory Aerodynamic System (PAS; KayPentax, Lincoln Park, NJ) was simultaneously recorded during productions of the expert SLP. Nasal airflow was captured at 200 Hz, and the microphone embedded in the PAS was recorded at 22025 Hz.

Nasal Acceleration Energy Measure

The NAEM was calculated for every production of /pɑ/. Microphone signals were first examined using Praat software (Boersma & Weenink, 2015) by one of two trained technicians. The microphone signal was viewed in Praat both as a time waveform and as a spectrogram. The burst of each /p/ in the syllable strings was located and noted. If the burst was not apparent in either the time waveform or spectrogram, or if the hold period was shorter than 25 ms, the /p/ was not included in the NAEM (see Figure 1). Next, custom MATLAB (MathWorks, Natick, MA) software was used to calculate the power density of the nasal accelerometer

Figure 1. Signals collected from a 6-year-old male child with a repaired submucosal cleft palate, while saying /pɑpɑ/ (roughly 600-ms duration). The top (marked NA) shows the time waveform of the nasal accelerometer signal; the middle shows the time waveform of the microphone signal; the bottom panel shows the spectrogram (0–5000 Hz on the *y*-axis) of the microphone signal. For the first /pɑ/, the burst is labeled with an arrow. Twenty-five milliseconds of the accelerometer signal during the hold period is selected and used for the nasal acceleration energy measure (see left of burst in accelerometer signal). For the second /pɑ/, no clear unvoiced hold period is detected. Thus, this instance of the utterance was not included in the NAEM calculation.

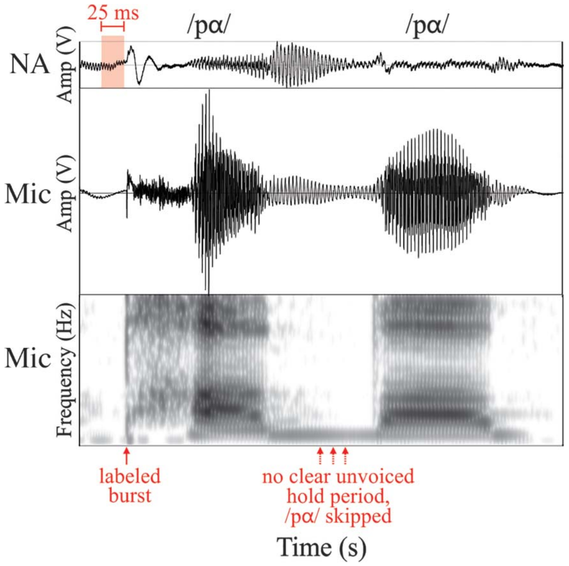
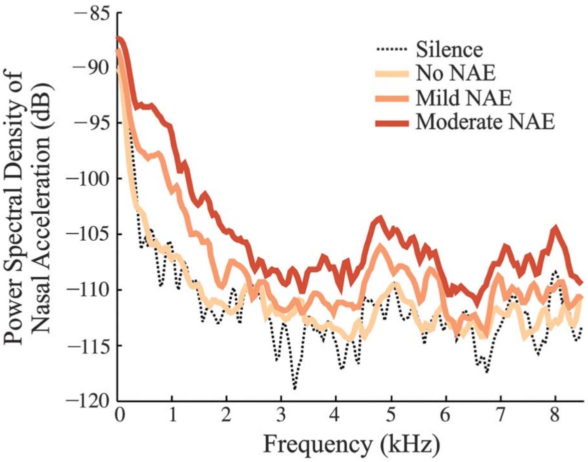


Figure 2. Average power spectra from nasal acceleration during NAE of different severities modeled by an SLP. Average spectrum for three sections of silence (quiet breathing) shown in black dotted line. Average power spectra for different levels of self-perceived NAE (none, mild, moderate) shown in shades of red; spectra are averaged over a series of /pɑpɑpɑpɑ/.



signal during the hold period. Although the reported average lengths of voiceless consonant hold periods range widely from 25 to 96 ms (Hixon, Weismer, & Hoit, 2014), we examined only the final 25 ms of the hold period; we chose this period of time both to ensure that the measure would not erroneously incorporate the preceding vowel, and because the end of the hold period is the point at which the most pressure is built up in the oral cavity. Welch's power spectral density was used with a window length of 10 ms and 5-ms overlap, in order to get power density estimates over the 25-ms period that were both reliable and sensitive. Power density estimates were obtained for each /pɑ/ and then averaged to get one power density estimate for a string of /pɑpɑpɑpɑ/. The final measure consisted of the power at 40–8500 Hz (due to accelerometer's mechanical resonance making readings above 8500 Hz unstable; Knowles Electronics, 2007) averaged over all three to 10 /pɑ/ hold periods, normalized by the average of the power at 40–8500 Hz from three periods of silence (quiet breathing). Average power spectra for different levels of NAE modeled by an SLP are shown in Figure 2. The power spectral density estimate across many frequencies increases as a function of increased self-perceived level of NAE. Note also that the spectrum for "no NAE" closely resembles the spectrum for periods of silence, indicating that during the hold periods of /pɑ/ tokens produced by a typical speaker, there was no air flowing through the nares to create additional power in the nasal accelerometer signal.

Perceptual Ratings of Pediatric Samples

To show initial validation of the proposed measure, a perceptual study was completed using the speech samples from the speakers with and without a history of cleft palate. Although there have been varied findings on whether inexperienced or experienced listeners yield the best interrater reliability (Brunnegård et al., 2012; Kreiman, Gerratt, Kempster, Erman, & Berke, 1993), we chose to use inexperienced listeners as they are more representative of the population as a whole. To improve reliability, it has been suggested that inexperienced listeners should receive training (Eadie & Baylor, 2006; Lee, Whitehill, & Ciocca, 2009) and/or practice rating voices with or without feedback on the accuracy or consistency of their ratings (Lee et al., 2009). Therefore, in this study, listeners were exposed to different severities of NAE and hypernasality during a training session, followed by practice without feedback in which they rated all samples for NAE while ignoring hypernasality prior to completing the final ratings.

Perceptual ratings were obtained from 14 inexperienced listeners (six men) with a mean age of 22.9 years. Listeners were required to pass a hearing screening (pure-tone air conduction at 25 dB HL at 250, 500, 750, 1000, 1500, 2000, 3000, 4000, 6000, and 8000 Hz). Listeners were provided with written explanations of NAE, hypernasality, and hyponasality, and were instructed to disregard any speech or resonance abnormalities in the samples that were not related to NAE. Listeners interacted with a graphical user interface that presented a training module followed by three direct magnitude estimation (DME) modules.

The training module allowed listeners to play four samples noted as containing typical speech and mild, moderate, and severe NAE, and two samples marked as mild and moderate hypernasality. These samples came from the same pediatric speakers as the experimental stimuli but were not the same samples used in the study; instead they consisted of alternate utterances such as /tatatata/ and /shishishi/. Samples were rated for severity by an expert SLP rater. Listeners were able to play samples as many times as they wished and were required to play each sample at least once before moving on to the DME modules.

Although DME is infeasible in a clinic setting, it can be used to capture reliable ratings of any type of variable (e.g., prothetic or metathetic; Whitehill, Lee, & Chun, 2002). In the DME modules, listeners used a standard DME with modulus procedure, in which raters listened to both a modulus sample and the sample to be rated. The modulus was the same for all trials and all listeners, and both the modulus and sample could be played as many times as the listener wanted. Listeners were required to listen to both the modulus and each sample on every trial and rate the sample as compared to the modulus. Listeners were instructed that the modulus had a score of 100. If the listener judged the sample to have NAE twice as severe as the modulus, he/she was instructed to rate the sample as 200; if the listener judged the sample to have NAE half as severe as the modulus, he/she was instructed to rate the sample as 50 (Whitehill et al., 2002). Listeners rated all randomly ordered samples three times, with a brief rest (1–5 mins, listener-directed) between each module.

Data Analyses

The nasal airflow and nasal acceleration from the expert SLP were time-aligned using the microphone signals (one associated with the nasal acceleration and one embedded in the PAS). Airflow was calculated in liters per second over the 25-ms period defined above (i.e., immediately preceding the burst of each /p/) and averaged over the utterance of seven /pɑ/ productions. The NAEM was calculated for each self-defined NAE severity level. A Pearson's correlation was used to compare the airflow to the NAEM.

The auditory-perceptual ratings completed during the first DME module were considered practice and not analyzed further. For each listener, perceptual ratings from the second and third DME modules were compared with Pearson's correlations to determine intrarater reliability. The geometric mean of each rater's second and third DME module ratings was calculated and used for the remainder of the analyses (Stevens, Kalikow, & Willemain, 1975; Whitehill et al., 2002; Zraick & Liss, 2000). Interrater reliability was calculated using the intraclass correlation coefficient, type 2k, with consistency (Shrout & Fleiss, 1979). The geometric mean of all raters was calculated for each sample, and a Pearson's correlation was used to compare these perceptual ratings to the NAEM.

Results

Expert SLP Productions

Figure 3 shows the NAEM compared to average nasal airflow during the same hold period; both measures are averaged over one string of /pɑpɑpɑpɑpɑpɑpɑ/. The Pearson's correlation between amount of nasal airflow and the nasal acceleration energy measure was high ($r = .87$).

Perceptual Results

Intrarater reliability between the second and third DME modules was calculated for each rater, as were interrater reliabilities between each pair of raters. Intrarater reliabilities ranged from .26 to .87 ($M = .61$) before four raters were excluded for having intrarater reliabilities below .5. Two additional raters were excluded for low interrater reliability (mean correlation with all other raters <.25), resulting in 14 final raters. The final range of intrarater reliabilities was from .53 to .78 ($M = .66$). Normality of perceptual ratings was verified with the Kolmogorov–Smirnov test, and the intraclass correlation coefficient was .85, showing strong consistency amongst raters. The Pearson's correlation between perceptual ratings and the NAEM was moderate ($r = .49$). Comparisons between the perceptual results and the NAEM are shown in Figure 4.

Discussion

The NAEM we have developed uses recordings of nasal skin vibration to assess the severity of NAE both within and across speakers. The initial validation presented

Figure 3. Comparison between the mean airflow and the NAEM during NAE modeled by the expert SLP participant. Mean airflow and NAEM were both calculated over the hold period of each /p/ and averaged over a string of seven /pɑ/ productions. The self-perceived severity of the modeled NAE is shown in different shades of red.

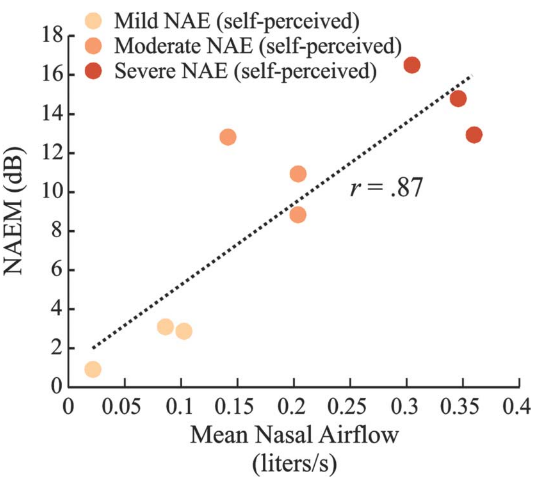
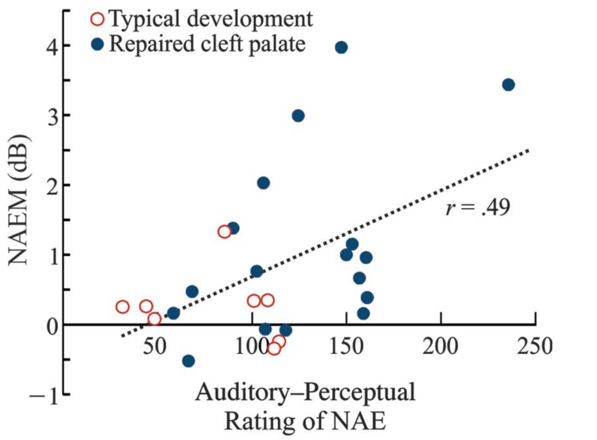


Figure 4. Comparison between auditory-perceptual ratings of NAE severity and the NAEM. Each point represents one speaker; speakers with repaired cleft palate are shown in filled blue circles, whereas speakers with no history of speech disorders are shown in empty red circles.



here shows a strong positive correlation with the nasal airflow produced at different NAE severities by an expert SLP (within speaker). Our validation study also shows a moderate positive correlation with auditory-perceptual ratings by inexperienced listeners (across speakers).

Comparison to Perceptual Ratings

Multiple factors may contribute to the variance seen when comparing NAEM with perceptual ratings. First, listeners are asked to rate NAEs that may be intrinsically inaudible. In addition, perceptual ratings are inherently subjective and often unreliable, despite being the most commonly used clinical assessment tool. Listener perception of NAE alone may be particularly unreliable as the percepts are subtle, and individuals with cleft palate may have concurrent hypernasal resonance and compensatory articulation (Brunnegård et al., 2009). In light of these constraints, the correlation with listener perception shown here is quite promising. Future work will both refine the NAEM and validate it against other quantitative measures, particularly to establish norms for speakers stratified by age and diagnosis, as needed.

Advantages and Disadvantages Compared to Nasalance and Aerodynamic Measures

Nasalance is a clinically available acoustic measure that provides a simple ratio of the acoustic energy from the nose and mouth. However, although nasalance can reliably distinguish between children with VPD and typically developing children, it is not sensitive to NAE, as high nasalance can result from acoustic energy due to either hypernasality or NAE. In addition, measuring nasalance requires an uncomfortable headset (Pratt & Hricisak, 1994) that has a low degree of cooperation in young children (van der Heijden, Hobbel, van der Laan, Korsten-Meijer, & Goorhuis-Brouwer, 2011) and is comparatively expensive ($3,000). These factors all suggest that nasalance may not be sensitive enough or feasible for clinical assessment of NAE.

Although aerodynamic measures may be more accurate than nasalance in determining the amount of air escaping through the VP port, these measures are used infrequently clinically (4.3% of clinicians involved in VPD management reported its use; Kummer et al., 2012). This is likely due to comparatively high cost (commercial systems are roughly $10,000), low client compliance, and, in some systems, low precision at the small amounts of emitted air to be expected from speakers with mild NAE.

On the other hand, the equipment originally proposed by Stevens et al. (1975) for VPD assessment and used here consists of very low-cost, commercially available components ($150), and is low weight, rugged, and easy to apply. Its decreased invasiveness as compared to nasal endoscopy and nasometry may lead to increased compliance; children with cleft palate and healthy controls show a preference for accelerometer-based instrumentation relative to nasometry instrumentation (Braden, Varghese, & Stepp, 2013).

Limitations and Future Directions

Most of the samples presented here were from speakers with no or mild NAE. Future studies will need to include more severe samples, both to further validate this measure and to provide listeners with a wider range of samples to better anchor their internal standards for different severity levels (Kreiman et al., 1993). Future studies will also involve further validation of the measure against airflow measures in many individuals with VPD, including individuals with craniofacial anomalies that include altered nasal structures, which may affect the mechanical vibrations sensed by the accelerometer. Airflow data collection is not well tolerated in children (nor is it used frequently in cleft palate/craniofacial clinics), but for research purposes may be tolerated by older children with VPD or adults with VPD due to hearing impairment or dysarthria.

Conclusion

We have presented a new quantitative measure of nasal emission utilizing nasal accelerometry in children with and without a history of cleft palate. This measure determined the power of the accelerometer signal in a wide range of frequencies during the hold period of a /p/ compared to during a period of quiet breathing. Nasal emissions modeled by an expert SLP were found to have a strong positive correlation (r = .87) between the quantitative measure and nasal airflow. The measure was then compared to perceptual ratings and found to have a moderate positive correlation (r = .49). Future directions will include further validation in this population and in other individuals with VPD.

Acknowledgments

This work was supported, in part, by grants to Cara Stepp from the Deborah Munroe Noonan Memorial Research Fund and Grant DC012651 from the National Institute on Deafness and Other Communication Disorders, as well as by a grant to Maia Braden through the Diane M. Bless Endowed Chair, Division of Otolaryngology–Head and Neck Surgery at University of Wisconsin–Madison. The authors would like to thank Victoria McKenna, Joseph Mendoza, and Carolyn Michener for their assistance.

References

Baylis, A., Chapman, K., Whitehill, T. L., & The Americleft Speech Group. (2015). Validity and reliability of visual analog scaling for assessment of hypernasality and audible nasal emission in children with repaired cleft palate. *The Cleft Palate–Craniofacial Journal, 52,* 660–670. doi:10.1597/14-040

Baylis, A. L., Munson, B., & Moller, K. T. (2011). Perceptions of audible nasal emission in speakers with cleft palate: A comparative study of listener judgments. *The Cleft Palate–Craniofacial Journal, 48,* 399–411. doi:10.1597/09-201

Boersma, P., & Weenink, D. (2015). *Praat: Doing phonetics by computer* (Version 5.4). Retrieved from http://www.praat.org/

Braden, M. N., Varghese, L. A., & Stepp, C. E. (2013). *Application of normalized nasal acceleration to children with and without cleft palate*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Chicago, IL.

Brancamp, T. U., Lewis, K. E., & Watterson, T. (2010). The relationship between nasalance scores and nasality ratings obtained with equal appearing interval and direct magnitude estimation scaling methods. *The Cleft Palate–Craniofacial Journal, 47,* 631–637.

Brunnegård, K., Lohmander, A., & van Doorn, J. (2009). Untrained listeners' ratings of speech disorders in a group with cleft palate: A comparison with speech and language pathologists' ratings. *International Journal of Language & Communication Disorders, 44,* 656–674. doi:10.1080/13682820802295203

Brunnegård, K., Lohmander, A., & van Doorn, J. (2012). Comparison between perceptual assessments of nasality and nasalance scores. *International Journal of Language & Communication Disorders, 47,* 556–566. doi:10.1111/j.1460-6984.2012.00165.x

Bunton, K., Hoit, J. D., & Gallagher, K. (2011). A simple technique for determining velopharyngeal status during speech production. *Seminars in Speech and Language, 32,* 69–80.

Dalston, R. M., Warren, D. W., & Dalston, E. T. (1991). A preliminary investigation concerning the use of nasometry in identifying patients with hyponasality and/or nasal airway impairment. *Journal of Speech and Hearing Research, 34,* 11–18.

Dotevall, H., Lohmander-Agerskov, A., Ejnell, H., & Bake, B. (2002). Perceptual evaluation of speech and velopharyngeal function in children with and without cleft palate and the relationship to nasal airflow patterns. *The Cleft Palate–Craniofacial Journal, 39,* 409–424.

Eadie, T. L., & Baylor, C. R. (2006). The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *Journal of Voice, 20,* 527–544.

Hixon, T. J., Weismer, G., & Hoit, J. D. (2014). *Preclinical speech science: Anatomy, physiology, acoustics, and perception* (2nd ed.). San Diego, CA: Plural.

John, A., Sell, D., Sweeney, T., Harding-Bell, A., & Williams, A. (2006). The cleft audit protocol for speech—augmented: A validated and reliable measure for auditing cleft speech. *The Cleft Palate–Craniofacial Journal, 43,* 272–288.

Karnell, M. P. (1995). Nasometric discrimination of hypernasality and turbulent nasal airflow. *The Cleft Palate–Craniofacial Journal, 32,* 145–148.

Knowles Electronics. (2007). *Knowles bu-21771 series accelerometer* [technical specifications]. Retrieved from http://www.knowles.com/eng/content/download/3058/35824/version/3/file/BU-21771-000.pdf

Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. *Journal of Speech and Hearing Research, 36,* 21–40.

Kummer, A. W. (2008). *Cleft palate and craniofacial anomalies: Effects on speech and resonance* (2nd ed.). Clifton Park, NY: Thomson Delmar Learning.

Kummer, A. W., Briggs, M., & Lee, L. (2003). The relationship between the characteristics of speech and velopharyngeal gap size. *The Cleft Palate–Craniofacial Journal, 40,* 590–596.

Kummer, A. W., Clark, S. L., Redle, E. E., Thomsen, L. L., & Billmire, D. A. (2012). Current practice in assessing and reporting speech outcomes of cleft palate and velopharyngeal surgery: A survey of cleft palate/craniofacial professionals. *The Cleft Palate–Craniofacial Journal, 49,* 146–152. doi:10.1597/10-285

Kummer, A. W., Curtis, C., Wiggs, M., Lee, L., & Strife, J. L. (1992). Comparison of velopharyngeal gap size in patients with hypernasality, hypernasality and nasal emission, or nasal turbulence (rustle) as the primary speech characteristic. *The Cleft Palate–Craniofacial Journal, 29,* 152–156.

Laczi, E., Sussman, J. E., Stathopoulos, E. T., & Huber, J. (2005). Perceptual evaluation of hypernasality compared to HONC measures: The role of experience. *The Cleft Palate–Craniofacial Journal, 42,* 202–211. doi:10.1597/03-011.1

Lee, A., Whitehill, T. L., & Ciocca, V. (2009). Effect of listener training on perceptual judgement of hypernasality. *Clinical Linguistics & Phonetics, 23,* 319–334.

MacKay, I. R., & Kummer, A. W. (1994). *Simplified nasometric assessment procedures*. Lincoln Park, NJ: Kay Elemetrics.

McWilliams, B., & Philips, B. (1979). *Velopharyngeal incompetence: Audio seminars in speech pathology*. Philadelphia, PA: Saunders.

Mra, Z., Sussman, J. E., & Fenwick, J. (1998). HONC measures in 4- to 6-year-old children. Horii oral nasal coupling index. *The Cleft Palate–Craniofacial Journal, 35,* 408–414. doi:10.1597/1545-1569(1998)035<0408:HMITYO>2.3.CO;2

Pratt, S. R., & Hricisak, I. (1994). Commercially available computer-based speech feedback systems. *Journal of the Academy of Rehabilitative Audiology, XXVII,* 89–106.

Redenbaugh, M. A., & Reich, A. R. (1985). Correspondence between an accelerometric nasal/voice amplitude ratio and listeners' direct magnitude estimations of hypernasality. *Journal of Speech and Hearing Disorders, 28,* 273–281.

Rintala, A., & Haapanen, M.-L. (1995). The correlation between training and skill of the surgeon and reoperation rate for persistent cleft palate speech. *British Journal of Oral and Maxillofacial Surgery, 33,* 295–298.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Stevens, K. N., Kalikow, D. N., & Willemain, T. R. (1975). A miniature accelerometer for detecting glottal waveforms and nasalization. *Journal of Speech and Hearing Disorders, 18,* 594–599.

van der Heijden, P., Hobbel, H. H., van der Laan, B. F., Korsten-Meijer, A. G., & Goorhuis-Brouwer, S. M. (2011). Nasometry cooperation in children 4–6 years of age. *International Journal of Pediatric Otorhinolaryngology, 75,* 627–630. doi:10.1016/j.ijporl.2011.01.035

Whitehill, T. L., Lee, A. S., & Chun, J. C. (2002). Direct magnitude estimation and interval scaling of hypernasality. *Journal of Speech, Language, and Hearing Research, 45,* 80–88.

Witt, P. D., Wahlen, J. C., Marsh, J. L., Grames, L. M., & Pilgram, T. K. (1998). The effect of surgeon experience on velopharyngeal functional outcome following palatoplasty: Is there a learning curve? *Plastic and Reconstructive Surgery, 102,* 1375–1384.

Wyatt, R., Sell, D., Russell, J., Harding, A., Harland, K., & Albery, L. (1996). Cleft palate speech dissected: A review of current knowledge and analysis. *British Journal of Plastic Surgery, 49,* 143–149.

Zraick, R. I., & Liss, J. M. (2000). A comparison of equal-appearing interval scaling and direct magnitude estimation of nasal voice quality. *Journal of Speech, Language, and Hearing Research, 43,* 979–988.