Aerodynamic Validation of Perceptually-Based Breath Group Determination and Temporal Breath Group Structure Analysis in Taiwanese Adolescents with Prelingual Severe to Profound Hearing Impairment

Wei-Chun Che[1], Yu-Tsai Wang[2], Hsiu-Jung Lu[2]

[1]Department of Physical and Medical Rehabilitation, National Taiwan University, Taipei, Taiwan
[2]School of Dentistry, National Yang-Ming University, Taipei, Taiwan
b88102080@ntu.edu.tw, yutsaiwang@ym.edu.tw, hsiujung_lu@hotmail.com

ABSTRACT

This study reported the reliability and validity of perceptually determined inspiratory loci and temporal breath group structure between 20 young Taiwanese adolescents with prelingual severe to profound hearing impairment (HI) and 20 age-gender-education-matched hearing controls (HC). The reliability and validity of perceptual judgment of inspiratory loci were considered satisfactory for both groups, although the HI group exhibited more error rate than the HC group. Furthermore, compared to the HC group, the HI group had more inappropriate inspiratory loci and speech breathing frequencies, longer inter-breath-group pause, but comparable breath group duration.

Keywords: breath group, aerodynamic analysis, validity, hearing impairment, acoustic analysis.

1. INTRODUCTION

Most people with prelingual severe to profound hearing-impairment have problems in the speech production subsystems. The past studies have emphasized on the phonation and articulation systems. Few studies, however, have focused on the speech breathing function and its impact on prosodic features in the HI. Breath group analysis can serve as a valid unit for prosodic analysis. Breath group analysis not only relates to speech intelligibility and naturalness, but also provides prosodic profile of speech. Investigations on the breath group structure in the HI population would be beneficial for clinical assessment and treatment.

The inspiratory loci of breath groups are usually determined by the following three methods: (a) use inspiratory signals recorded by instruments, e.g., Respirtrace [1], pneumotachography, or magnetometer [2,3]; (b) set up a fixed pause duration as a standard for inspiratory loci. For example, Yunusova, Weismer, Kent and Rusche [4] regarded pauses longer than 150 ms as inspiratory loci; Campbell and Dollahan [5] used 300 ms; and Walker, Archibald, Cherniak, and Fish [6] used 250 ms; (c) rely on perceptual judgment [7]. The third method is mostly used clinically. Major clues for perceptual judgment include audible inspiration and other acoustic cues, such as pauses at major syntactic boundaries, final phrase lengthening, and f_o downtrend. Past research has shown that inter-rater and intra-rater reliability for perceptual judgment of inspiratory loci for dysarthric speech and normal speech are satisfying [7]. However, for the HI population, none related data about the validity and reliability of perceptually determined breath group has been reported by now.

Thus our study aims to answer the following questions: (1) what are the reliability and validity of perceptually determined inspiratory loci in HI speakers? (2) Is there a difference in the temporal breath group characteristics between people with normal hearing and with prelingual severe to profound hearing impairment?

2. METHOD

While holding a circumferentially vented mask connected to a pneumotach, each participant read three short passages and the speech signal was recorded simultaneously. Inspiratory loci determined auditory-perceptually by three judges were used to test inter-judge reliability. Airflow traces were used to locate exact loci of inspiration on each speech signal to perform a validity test. Breath groups based on exact inspiratory loci were used for temporal breath group structure analysis.

2.1. Participants

HI participants were 20 Taiwanese adolescents (9 females and 11 males with age average and standard deviation of 18.7 and 2.4 years) with hearing threshold greater than 70 dB averaging over 500 Hz, 1000Hz, and 2000Hz on their better ear. HC participants were 20 age-gender-education-matched young adults and passed a bilateral pure tone screening at 25 dB HL for 500 Hz, 1000Hz, and 2000Hz. All participants were native Mandarin speakers without a history of respiratory, neurological, cognitive, or craniofacial impairment.

2.2. Procedures

The participants held a circumferentially vented mask (Glottal Enterprises MA-1L) to their face, which was connected to Aerophone II (Kay PENTAX, Model 6800, AP-2) through an air-flow transducer. The system can distinguish between inspiratory and expiratory flow direction. A head-mounted microphone (Crown CM311A) was placed approximately 2-4 cm away from the vented mask. Participants were asked to read the three selected passages at a comfortable speaking rate and loudness. The audio signal was recorded using a digital audio recorder (Marantz PMD 660) at 48 kHz sampling rate and 16-bit quantization level.

2.3. Reliability

The audio signals were used to judge inspiratory loci through listening by three trained native Mandarin speakers. The speech samples were orthographically transcribed by an experimenter. All punctuations of the orthographic transcription text were removed in order to guard against judges' reanalysis of breath groups based on the sequence of words in the transcript. The space between words in the prepared transcript text was three times standard spacing in order to reduce the effects of word sequence. The judges were asked to listen to the speech samples and to mark on the transcription sheets at which inspiration occurs based on an audible inspiration or various acoustic cues, such as longer pause duration, f_0 declination, and longer phrase-final duration. Judges were allowed to listen to the tape repeatedly until they were confident in their determination on the breath group locus. The agreement of perceptual determination of inspiratory loci was compared across the three judges to gauge the reliability.

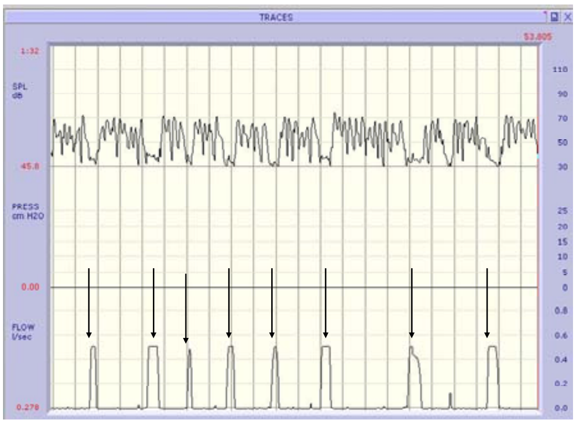
2.4. Validity

An experimenter marked all the onsets of an inspiration, as indicated by a peak on the air-flow transducer trace (i.e. inhalation), as a key (Figure 1). The perception-based loci of inspiration on which at least two out of three judges agreed (PERCEPT) were compared to the exact loci of inspiration determined by the aerodynamic signals (EXACT). For each task, the following measures were counted first: total number of EXACT, PERCEPT, loci where judges noted an inspiration that did occur (CORRECT), loci where judges noted an inspiration that did not occur (FALSE(+) = PERCEPT - CORRECT), and loci where judges missed an inspiration that did occur (FALSE(-) = EXACT - CORRECT). Then validity, error rate (ER), and missed rate (MR) were calculated according to the following formula:

$$Validity = \frac{CORRECT}{EXACT};$$

$$ER = \frac{FALSE(+)}{EXACT}; \quad MR = \frac{FALSE(-)}{EXACT}$$

Figure 1: The arrows indicated inspiratory loci based on aerodynamic signal for part of a passage read by a participant with profound hearing impairment.



2.5. Appropriateness of inspiratory loci

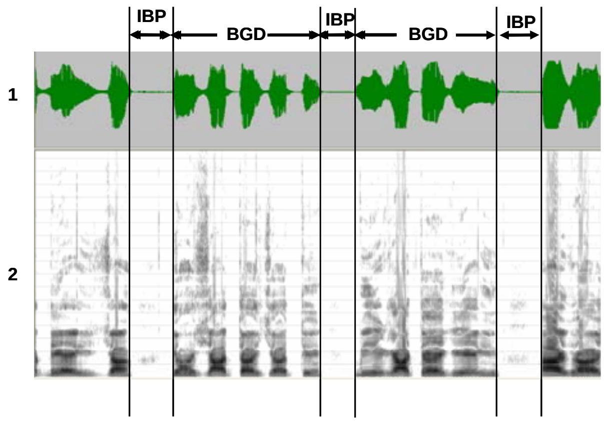
The appropriateness of inspiratory loci for the speech samples was determined on the basis of the rules given by Henderson and Goldman-Eisler [8] and Li and Thompson [9]. Appropriate syntactic boundaries for inspiration were defined as the location before noun, verb, adverbial phrases or other phrases, or where punctuation points, like comma, periods, occur. Inspiratory pauses within the phrases were considered syntactically inappropriate locations. The appropriateness of inspiratory loci was determined by a judge with training in linguistics.

2.6. Temporal breath group structure

The breath groups segmented by the actual inspiratory loci were used to acoustically analyze the temporal structure of breath group for the two groups. Acoustic measures were performed using

CSpeech software, TF-32 [10] (Milenkovic, 2001). Figure 2 shows loci of inspiration and measures of breath group structure based on acoustic and aerodynamic signals. As shown in figure 2, inspiratory loci were used to segment acoustic signals into BGD and inter-breath-group pause (IBP). BGD was defined as groups of speech events on a single breath [11], and was measured from the onset to the offset of acoustic events of a breath group based on the acoustic waveform and spectrogram. IBP was measured between BGDs. The first and last BGDs for each passage reading sample were excluded to avoid the influence of speech initiation and termination. The number of the breath groups and the time used to complete the tasks were first counted for each participant.

Figure 2: Acoustic signal for part of a passage read by a participant with profound hearing impairment. Panels 1 and 2 represent the waveform and wide-band spectrogram, respectively. BGD = breath group duration; IBP = inter-breath-group pause



3. RESULTS

3.1. Reliability

The inter-judge reliability consistent across at least two of the three judges was 1024/1168 (0.88) for the HI group. For the HC group, the inter-judge reliability consistent across at least two of the three judges was 587/682 (0.86). The reliability of perceptually-based breath group determination was comparable between the two groups.

3.2. Validity

For the HI group, the total number of inspirations based on aerodynamic signals (EXACT) was 1162. Comparing these instrumentally-determined events with the 1024 perceptually determined events on which at least 2 of the 3 judges agreed (PERCEPT), the results were as follows: There were 964 CORRECT, 60 FALSE(+), and 198 FALSE(-) events. This gives a validity of 964/1162 (83%), a MR of 198/1162 (17%), and an ER of 60/1162 (5%).

For the HC group, the EXACT was 709. Comparing these instrumentally-determined events with the 587 PERCEPT, the results were as follows: There were 579 CORRECT, 8 FALSE(+) , and 130 FALSE(-) events. This gives a validity of 579/709 (82%), a MR of 130/709 (18%), and an ER of 8/709 (1%).

The validity and MR of perceptually-based inspiratory loci determination was comparable between the HI and HC groups. But the HI group exhibited 5 times error rate on perceptual judgment of inspiratory loci than the HC group, i.e the rate of missed inspirations was comparable for both groups, but the HI group exhibited a much higher rate of noting inspirations where none occurred.

3.3. Appropriateness of inspiratory loci

The inappropriate inspiratory loci for the HI and HC groups were 94 out of 1162 (8.08%) and 5 out of 709 (0.7%), respectively. The number of inappropriate inspiratory loci for the HI group was more than 10 times than that for the HC group.

3.4. Breath group temporal structure

Compared to the HC group, the HI group needed more respiratory cycles and speaking time to complete the tasks as indicated in Table 1. With respect to the BGD, two groups were comparable. However, the average IBP for the HI group was much longer than that for the HC group. The variation of IBP for the HI group was also larger than that for the HC group.

TABLE 1 . Group means, standard deviations, maxiimum and minimum of all the parameters for the temporal breath group structures .

	HI			HC		
	mean (SD)	min.	max.	mean (SD)	min.	max.
# of BG	55.1 (17.8)	35	99	32.5 (9.5)	18	53
ST (s)	158.7 (38.0)	95.4	252.1	101.4 (10.5)	88.0	120.1
BGD (s)	2.11 (0.44)	0.18	7.42	2.16 (0.59)	0.38	7.82
IBP (ms)	561 (127)	200	1984	458 (95)	198	1472

Note. #BG = number of breath groups to complete the speech tasks
ST = total speaking time to complete the speech tasks

4. DISCUSSION

Speech breathing in the HI speakers might be impaired due to impaired auditory feedback [12], but the perception of inspiratory loci for the HI speakers is still discernible in tonal language. The considered satisfactory reliability and validity of perceptually-based breath group determination indicated that the clinical practice of perceiving inspiratory loci might be practical. Although the validity of perceptual judgment for both groups was comparable, it is more likely to falsely perceive inspiratory loci for the HI group than for the HC group. Preliminary examination on the data revealed some possible factors that affect the perceptual judgment of inspiratory loci for the HI group: longer pauses within the breath group, f_0 downtrend due to glottalization, and the improper usage of airflow causing noises in speech production.

The minimum IBP (200ms) for both groups in the present study indicated that using certain fixed pause duration as a standard for inspiratory loci might render some false positives or false negatives, such as using 150 ms in Yunusova et al. [4] or 300 ms in Campbell and Dollaghan[5]. Its significance needs to be verified by further scrutiny on the pause duration within the breath groups.

Moreover, the HI group exhibited more inappropriate inspiratory loci than the HC group. The results were comparable with that of research in English-speaking HI speakers [13]. The poor respiratory airflow control and coordination among the speech subsystems in the HI group might contribute to compulsory pauses for inspiration even at syntactically inappropriate locations. A greater proportion of inappropriate inspiratory loci might affect speech intelligibility [14]. Therefore, inappropriate inspiratory pauses might be one of the possible reasons for the reduced intelligibility in the HI group.

It is important to note that certain features of temporal breath group structure was preserved in the HI group, but others were not. BGD was generally intact in these participants; however, speech breathing cycles and average IBP were more frequent and longer. Based on the results given, it seems highly likely that speech rate is lower for the HI group.

The longer average IBP for the HI group than the HC group implied possible factors caused by the long-term severe to profound hearing loss, such as wastage of lung air before speech initiation [13], more inappropriate inspiratory loci, and the inefficient coordination among respiratory, phonatory, and articulatory subsystems of speech production [13]. In addition, the non-inspiratory pause, defined as the difference between inter-breath-group pause and inspiratory duration, might be a good index of the inefficient coordination among subsystems of speech production. Further studies on the performance of this index are needed to clarify this issue.

5. CONCLUSION

The reliability and validity of perceptual judgment of inspiratory loci were considered satisfactory for both the HI and HC groups in the present study. Compared to the HC group, the HI group exhibited similar BGD but significantly more frequent speech breathing cycle and longer IBP.

6. REFERENCES

[1] Winkworth, A.L., Davis, P.J., Ellis, E., Adams, R.D. 1994. Variability and consistency in speech breathing during reading: lung volumes, speech intensity, and linguistic factors. *Journal of Speech and Hearing Research*, 37, 535-556.

[2] Hoit, J.D., Hixon, T.J., Watson, P.J., & Morgan, W.J. 1990. Speech breathing in children and adolescents. *Journal of Speech and Hearing Research,* 33, 51-69.

[3] Bunton, K. 2005. Patterns of lung volume use during an extemporaneous speech task in persons with Parkinson disease. *Journal of Communication Disorders,* 38, 331-348.

[4] Yunusova, Y., Weismer, G., Kent, R.D., Rusche, N.M. 2005. Breath-Group Intelligibility in Dysarthria: Characteristics and Underlying Correlates. *Journal of Speech, Language, and Hearing Research,* 48, 1294-1310.

[5] Campbell, T.F. Dollaghan, C.A. 1995. Speaking rate, articulatory speed, and linguistic processing in children and adolescents with severe traumatic brain injury. *Journal of Speech and Hearing Research,* 38, 864-875.

[6] Walker, J.F., Archibald, L.M., Cherniak, S.R., Fish, V.G. 1992. Articulation rate in 3- and 5-year-old children. *Journal of Speech and Hearing Research,* 35, 4-13.

[7] Wang, Y.T., Kent, R.D., Duffy, J.R., Thomas, J.E. 2005. Dysarthria in traumatic brain injury: a breath group and intonational analysis. *Folia phoniatrica et logopaedica,* 57, 59-89.

[8] Henderson, A., Goldman-Eisler, F., Skarbek, A. 1965. Temporal patterns of cognitive activity and breath control in speech. *Language & Speech*, 8, 236-42.

[9] Li, C.N., Thompson, S.A. 1989. *Mandarin chinese: A functional reference grammar* University of California.

[10] Milenkovic, P. 2001. Time-frequency analysis for 32-bit windows [Computer software]. Madison: Wisconsin.

[11] Kent, R.D., Read, C. 2002. *The Acoustic Analysis of Speech.* San Diego, California: Singular.

[12] Tye-Murray, N. 1992. Ariculatory organizational strategies and the role of audition. *Volta Review*, 94, 243-260.

[13] Forner, L.L., Hixon, T.J. 1977. Respiratory kinematics in profoundly hearing-impaired speakers. *Journal of Speech and Hearing Research,* 20, 373-408.

[14] Hammen, V.L., Yorkston, K.M. 1994. Respiratory patterning and variability in dysarthric speech. *Journal of Medical Speech-Language Pathology*, 2, 253-261.