Comparison Study of Vocal Attack Time in Patients With Cleft Palate With and Without Glottal Stop in Mandarin

*Zhen Ren, *Xia Zhou, *Lian Ma, and †Jiangping Kong, *†*Beijing, China*

Summary: Glottal stop (GS) is a typical compensatory articulation, which has a great impact on speech intelligibility in patients with cleft palate. It is usually detected by perceptual analysis. The aim of this study is to investigate the utility of vocal attack time (VAT) values in patients with cleft palate with and without GS, when unaspirated monosyllables are articulated in Mandarin, by using electroglottography. Unaspirated monosyllables /pa/ /pi/ /pu/ /ta/ /ti/ /tu/ /ka/ /ki/ /ku/ with tone one were analyzed. A total of 575 tokens were obtained from 42 patients with cleft palate, divided into a GS category (n = 312 tokens) and a nonglottal stop (NGS) category (n = 263 tokens), as assessed perceptually by three judges. Sound pressure and electroglottography recordings were also obtained from these tokens. The time lag of the cross-correlation function was used to gain VAT values. The results showed that the mean VAT values of tokens from the GS category (−0.25 ms) was significantly shorter than that of tokens in the NGS category (3.19 ms) (t = 7.326, $P < 0.001$). The results also showed that there was no significant difference in VAT values between the different combined monosyllables both in GS and in NGS group. The conclusion that can be drawn from this study is that the VAT value was sensitively decreased in cleft palate Mandarin speakers with GS comparing to those without GS.
Key Words: Cleft palate−Glottal stop−Mandarin−Vocal attack time−Electroglottography.

INTRODUCTION

Cleft palate and/or lip are recognized as a common congenital craniofacial malformation. Compensatory articulation disorders are generally the most common speech abnormalities, which also include abnormal resonance (hypernasality and/or hyponasality), voice disorder, and nasal emission for children with clefts.[1] As the most frequent compensatory articulation error, glottal stop (GS), which is noticeable, distracting, and socially stigmatizing, has several impacts on speech and language leading to negative effects on social communication and the quality of life.[2,3]

According to the paper reviewed by Warren,[4] a GS is produced when air pressure loss occurs due to velopharyngeal insufficiency. A compensation mechanism works to regulate pressure in the vocal tract by adducting the vocal folds abruptly. In Mandarin, a GS is the substitution for a plosive consonant in cleft palate speech. While rarely observed in normal speech in Mandarin, GS exists in other languages like English, Spanish, Persian, and even as a phoneme in dialect of Chinese.

Many of the studies documenting speech production in cleft lip and palate have been based on perceptual analysis of speech data, using either phonetic transcription or some sort of ratings scale assessing general intelligibility and/or the presence of features such as hypernasality, nasal escape, or compensatory articulations.[5−8] Perceptual analyses, however skilled, are not always able to identify speech production behaviors accurately.[9,10] However, the primary concerns with perceptual analyses are inter/intrarater reliability than the accuracy itself. Instrumental evaluation is necessary for detecting GS in patient with cleft palate due to the inaccuracy of perceptual analyses and the lack of experienced experts in China.

Fabre described the application of electroglottography (EGG), which was used to investigate laryngeal function.[11] As the details of original and improved electroglottographic methods became more widely available, it has been quickly adopted to examine normal glottal behaviors. EGG was applied to patients with congenital cleft palate, then some researchers investigated hoarseness in patients with cleft palate through EGG in recent years.[12−15] Some studies also indicated that EGG was a reliable instrument to reveal the acoustic measures of GS in different languages, such as English, Spanish, and some dialects in Chinese.[16−19]

Baken and Orlikoff[20] developed a simple and noninvasive method to measure the duration of vocal onset through obtaining sound pressure (SP) and EGG signals simultaneously, which is related to characteristics of attack gesture. The intersignal lag, termed vocal attack time (VAT), can be automatically and objectively estimated using a cross-correlation method and is, therefore, free of operator/investigator bias and imprecision.[21] The VAT measure was validated by Orlikoff et al[21] with high-speed video endoscopy, from which a digital kymogram was generated. In 2012, Roark et al[22] proposed a criterion, a figure of merit, which assessed a critical assumption of vocal startup on which the VAT measure is based and therefore represents integrity of the derived measure. Watson et al,[23,24] Ma et al,[25] Lee,[26] Tam,[27] and Zhang et al[28,29] reported normative VAT values for English, Cantonese, and Mandarin speakers, respectively. Consequently, VAT can be taken to be the duration from the start of vocal cord oscillation to the instant for the first vocal cord contact and provides a useful index to indicate the prephonatory laryngeal adjustment.

To the best of our knowledge, no EGG and VAT measures have been applied to the GS in Mandarin patients with cleft

Accepted for publication April 27, 2018.
From the *Department of Oral and Maxillofacial Surgery, Peking University School and Hospital of Stomatology, Beijing, China; and the †Department of Chinese Language and Literature, Center for Chinese Linguistics, Peking University, Beijing, China.

Address correspondence and reprint requests to Lian Ma, Department of Oral and Maxillofacial Surgery, Peking University School and Hospital of Stomatology, 22 Zhongguancun Avenue South, Haidian District, Beijing 100081, China.

Address correspondence and reprint requests to Jiangping Kong Department of Chinese Language and Literature, Center for Chinese Linguistics, Peking University. No. 5 Yiheyuan Road, Haidian District, Beijing, China. E-mail addresses: lamaiana2015@163.com jpkong@pku.edu.cn
Journal of Voice, Vol. 33, No. 5, pp. 803.e15−803.e21
0892-1997

https://doi.org/10.1016/j.jvoice.2018.04.020

palate. In this study, EGG and VAT measures were applied for the first time to identify and to characterize the association of VAT and GS in Mandarin patients with cleft palate, which will bring great advantages to assess compensatory articulation objectively for the patients with cleft palate. VAT holds promise as a meaningful clinical indicator of the GS of patients with cleft palate. It is also likely to be useful in the exploration of the finer details of vocal physiology on a more theoretical plane.

MATERIALS AND METHODS

Subjects and instrumentation

This project was approved by the Human Research Ethics Committee of the Stomatology Hospital of Peking University (no. PKUSSIR201734025), and informed consent was obtained from the patients' legal representatives. Forty-two patients (30 males; 12 females) with cleft palate aged from 6 to 46 years (mean = 18.70) were recruited into the current study. The cleft type included unilateral cleft lip and palate, bilateral cleft lip and palate, isolated cleft palate, and submucous cleft palate. All of the patients were native Mandarin speakers. They were visiting the outpatient department of the Cleft Lip and Palate Clinic Center, Stomatology Hospital of Peking University for further consultation. The recording was accomplished in the sound-treated booth at the Cleft Lip and Palate Clinic Center, Stomatology Hospital of Peking University, where the background noise was below 45 dBA.

Before data collection, the skin overlying the thyroid cartilage was cleaned using an alcohol swab before placement of surface electrodes for the EGG (Model 6103, KayPENTAX, Montvale, NJ). Patients wore a unidirectional collar clip microphone (ECM-44B, SONY, Tokyo, Japan) for detection of the SP signal. The microphone was positioned approximately 5 cm in front of the patient's mouth, just off the midline. The EGG and SP signals were routed through an amplifier and digitizer (sound card model Sound Blaster X-Fi Surround 5.1 Pro, Creative Labs, Jurong East, Singapore) to the laptop computer (Lenovo, x220i, Lenovo, Beijing, China) through a sound console (Behringer XENYX502, MUSIC Group Macao Commercial Offshore Limited, Zhongshan, Guangdong, China). Signal acquisition was accomplished using *Adobe Audition 3.0 software* (Adobe Systems Incorporated, San Jose, CA, USA) set to a sampling rate of 44,100 Hz and 32-bit resolution/channel. Gain was adjusted to ensure adequate amplification and to avoid clipping while patients said the syllable at comfortable pitch and loudness. The EGG and SP signals were monitored on a two-channel oscilloscope throughout data collection.

METHOD

The stimuli /pi/ /pa/ /pu/ /ti/ /ta/ /tu/ /ki/ /ka/ /ku/ consisting of three different unaspirated consonants with three different vertex vowels were printed on A4 paper. The patient was instructed to read each item at a comfortable pitch, loudness, and rate, and repeat every monosyllable 3 times (9×3). All the monosyllables were pronounced at tone one by patients. They were specifically instructed that this was not a reaction time task and to wait approximately 2−3 seconds before reading each monosyllable. EGG and SP signals were recorded continuously until the patient completed all tasks. EGG and SP signals were segmented postrecording to isolate individual utterances and were stored as separate files.

The utterances were also assessed by three judges who had clinical experience related to speech and language in cleft lip and palate children for more than 20 years. The tokens were divided into two categories based on the experts' perceptual judgments the nonglottal stop (NGS) category included utterances assessed by all three judges as not including a GS; the GS category included all tokens assessed by all three judges as containing a GS. Tokens on which one or more of the judges disagreed were abandoned.

VAT values were extracted from the signal database using computer-based automated software developed by Roark et al.[30] The data analysis process consists of four components. The first component is signal verification that provided for audiovisual inspection of the raw SP and EGG signals. The second component is signal segmentation that automatically identified a 600-ms segment of the SP and EGG signals that was centered at the approximate time of vocal onset. During the F0-based frequency filtering and signal-modeling component, the instantaneous amplitudes and frequencies were gained of modeled SP and EGG, which form the bandpass filtered model. The last component is extraction of measures so that the time lag (VAT) can be determined by a cross-correlation method, which is free of operator bias and imprecision.

Data analysis

The inter-rater reliability of the speech therapists' evaluations was evaluated using Cohen's "weighted Kappa-coefficient". All Kappa values were calculated at a confidence level of 95%. To obtain intrarater reliability, 10% of the entire stimuli were judged a second time by the three judges 1 month after the initial data collection.

VAT values were categorized by perceptual analysis into NGS and GS group. The variance between VAT values and the two variables, monosyllable and perceptual analysis, were analyzed statistically. VAT values of nine kinds of monosyllable were analyzed by analysis of variance (ANOVA) in NGS and GS group. An independent *t* test was performed to explore the changes of VAT values between GS group and NGS group.

RESULTS

From the 1134 utterances, as 499 utterances were abandoned due to the disagreements in the perceptual analysis of the three judges, 635 VAT values (351 for GS group and 284 for NGS group) were available for instrumental analysis. The abandoned utterances included 265 utterances that were classified as NGS by 2 judges (and 1 judge as GS) while 234 utterances were classified as GS by 2 judges (and 1 judge as NGS). Following the method described by Roark et al,[22] application of the figure of merit < 0.75 criterion to the database of measures resulted in the rejection of 5.8% of the recorded tokens (37 of

TABLE 1.
Inter-rater Reliability of Judges

	Judge 1/2	Judge 2/3	Judge 1/3
Kappa	0.622	0.652	0.558

635). A further 23 VAT values that were beyond ±2 standard deviations from the mean were also excluded. Accordingly, this report is based on the analysis of the remaining 575 tokens.

The degree of inter- and intrarater reliability of experienced judges' findings was moderate to good, as seen in Tables 1 and 2.

Table 3 summarized descriptive statistics for VAT values for the complete sample and for subsamples stratified by category and monosyllable. The mean VAT value for all 575 tokens was 1.32 ms, with a standard deviation of 6.21 ms. The ANOVA test indicated that there was no significant difference in VAT values between the nine kinds of monosyllable within each of the NGS and GS categories ($P > 0.05$). Within each of the categories, there was no significant difference in VAT values between tokens which were all unaspirated monosyllables of the same vowel combined with different consonants (eg, /pa/ /ta/ /ka/) or of the same consonant combined with different vertex vowels (eg, /pa/ /pi/ /pu/). VAT mean values and standard deviations are summarized in Table 4 and Figure 1.

TABLE 2.
Intra-rater Reliability of Judges

	Judge 1	Judge 2	Judge 3
Intra-rater reliability	81%	80%	83%

While there was no significant difference between the nine kinds of monosyllable within categories, the mean VAT value of all tokens in the GS category (−0.25 ms) was significantly different from the VAT value of tokens in the NGS category (3.19 ms) by independent t test ($t\ (575) = 7.326$, $P < 0.001$). (Table 5)

DISCUSSION

Perceptual analysis is frequently applied in clinical practice and in scientific evaluation by therapists because of its availability and easy applicability.[31] Regarding the features of articulation, the perceptual analysis results proved with near certainty that experienced judges' inter-rater reliability is moderate to good according to our results in Tables 1 and 2. Some researchers have indicated that there is a significant correlation between intelligibility and GS substitution—a typical compensatory articulation.[8] Unintelligibility in individuals with cleft palate is influenced by many factors, including the severity of the speech disorder, the speech stimuli used, the speech sample, the rate of speech, and the measurement task employed, any of which may decrease the accuracy of perceptual analysis.[32] Hikita et al[33] demonstrated that the reaction time for the perception

TABLE 3.
Descriptive Statistics of VAT Value for All Tokens and for Selected Subgroups Categorized by Monosyllables

		VAT			
Group	Number of Token	Mean (SD)	95% Confidence Interval	Minimum	Maximum
All tokens	575	1.32 (6.21)	0.80−1.82	−21.93	22.49
Tasks of NGS group					
/pa/	28	3.80 (2.84)	2.70−4.90	−3.26	8.73
/pi/	30	3.20 (3.79)	0.69−5.34	−5.83	9.43
/pu/	30	2.98 (3.28)	1.77−4.21	−4.47	13.15
/ta/	29	3.39 (2.80)	2.32−4.46	−2.59	9.55
/ti/	31	3.00 (2.47)	1.45−4.15	−6.89	9.86
/tu/	28	4.49 (5.22)	2.04−3.96	−2.86	7.30
/ka/	31	3.75 (3.36)	2.51−4.99	−4.49	11.70
/ki/	25	2.70 (3.81)	1.13−4.28	−5.97	11.43
/ku/	31	3.06 (3.45)	1.79−4.33	−5.60	11.56
Tasks of GS group					
/pa/	29	−1.72 (6.51)	−4.20 to 0.75	−16.25	10.41
/pi/	40	−0.49 (8.92)	−3.34 to 2.36	−15.24	19.77
/pu/	29	3.76 (8.92)	0.36−7.15	−21.45	21.54
/ta/	30	−1.48 (8.50)	−4.66 to 1.69	−13.81	22.44
/ti/	51	−0.72 (7.39)	−2.79 to 1.36	−21.92	12.54
/tu/	37	−2.26 (6.72)	−4.50 to −0.02	−12.20	16.83
/ka/	32	0.18 (6.26)	−2.07 to 2.44	−11.52	14.65
/ki/	33	1.16 (6.10)	−1.00 to 3.32	−14.60	13.58
/ku/	31	0.03 (6.83)	−2.48 to 2.53	−13.86	14.78

Abbreviation: SD, standard deviation.

TABLE 4.
Descriptive Statistics of VAT for All Tokens of Selected Subgroups Categorized by Different Vowels and Consonants

Group	NGS		GS	
Number	n = 263		n = 312	
Monosyllable	Mean VAT	SD VAT	Mean VAT	SD VAT
Same vowel with different consonants				
/pa/ /ta/ /ka/	3.64	3.00	−0.97	7.12
/pi/ /ti/ /ki/	2.91	3.72	−0.15	7.60
/pu/ /tu/ /ku/	3.01	3.08	0.27	7.80
Same consonant with different vowels				
/pa/ /pi/ /pu/	3.32	3.32	0.40	8.50
/ta/ /ti/ /tu/	3.06	3.02	−1.39	7.45
/ka/ /ki/ /ku/	3.20	3.51	0.47	6.35

Abbreviation: SD, standard deviation.

of GS is significantly longer than that for normal articulation, and it may require more auditory attention and more complex auditory processing than the perception of normal sounds. According to the current consensus, multiple-listener judgments are preferable to single-listener judgments even though intra-listener reliability has been demonstrated. Besides, training increases reliability.[7] In this study, although it proved the reliability of three skilled speech therapists' perceptual analysis detecting GS, the utterances with disagreement among three judges were abandoned to get as reliable results as possible for instrumental analysis.

GS can substitute for any high-pressure consonant in patients with cleft palate. For example, Chen et al,[34] using perceptual analysis, indicated that the frequency of GS in place of unaspirated plosive was the highest in Mandarin consonants from 12 patients with GS. In this study, three unaspirated plosive consonants at three places of articulation (bilabial, alveolar, and velar) were combined with three different vertex vowels as task monosyllables. According to the ANOVA test, there was no significant difference between the nine kinds of monosyllable. Watson et al[24] revealed a significant effect between aspirated ("hallways") and unaspirated ("always") task using VAT. In our research, the effects showed no difference across the three tasks, which were all unaspirated monosyllables of same vowel combined with different consonants (eg, /pa/ /ta/ /ka/) (Table 4). The mean VAT value of tasks (/pa/ /ta/ /ka/) in the NGS category (3.64 ms) was greater than the mean VAT value of "always" tasks (1.83 ms) and sustained /ɑ/ tasks (1.89 ms) reported by Watson et al.[24] The exact cause of the apparent difference needs further study. However, the most likely basis is the abnormal structure of articulators in patients with cleft palate, even after primary surgical closure of the cleft palate. The second basis affecting the VAT was due to the difference in methodology, because the English speakers produced a semantically empty monosyllable, whereas the Mandarin utterances were unaspirated monosyllable (consonant + vowel, C + V).[35]

Zhang et al[28] indicated that the three vertex Chinese vowels display their intrinsic VAT in such a pattern: /u/ > /i/ > /A/. The VAT values of the same vowel combined with different consonants (eg, /pa/ /ta/ /ka/) also showed an order in the NGS and GS categories. The mean GS VAT values displayed the same order as mentioned before: /pu/ /tu/ /ku/ > /pi/ /ti/ /ki/ > /pa/ /ta/ /ka/, while the NGS values displayed a different order : /pa/ /ta/ /ka/ > /pu/ /tu/ /ku/ > /pi/ /ti/ /ki/.

It was reported by Orlikoff et al[21] through five subjects (three men and two women) that the "breathy" onsets have positive VAT values (7.6−38.0 ms), the comfortable onset was typically shorter, ranging from −1.4 to 9.6 ms, the hard glottal attack was negative (ranging from −9.5 to −1.7 ms). This hard vocal attack pattern was described by Moore[36] based largely on an auditory perceptual categorization. Hirose[37] termed the attack referring to speech utterance consisting of C + V sequence, hard attack should be equivalent to the utterance initiated with GS.

The results here supported their findings. Table 5 indicated that the mean VAT value of all tokens in NGS group was positive (3.19 ms) and the mean VAT value in GS group was negative (−0.25 ms), suggesting that the mean VAT value tended to be related with perceptual analysis of GS in cleft palate.

Hirose[37] investigated the three types of vocal attack through electromyographic, characterizing that the hard glottal attack was the temporal pattern of adductor muscle activity. During the prephonatory period, the marked increase of adductor muscle activity related to the strong medial compression or constriction of the glottis before release. According to the physiological mechanism controlling abrupt glottal release after the period of constriction, the adductor activity showed a steep fall.

GS is a glottal sound like a hard vocal attack in patients with cleft palate, produced by the abrupt adduction and abduction of the vocal fold.[38] This process of compensatory behavior was exactly same process as mentioned before by Hirose. It occurs in substitution of the consonants that require higher intraoral pressure in patients with cleft palate with velopharyngeal insufficiency.

Some researchers investigated the GS or glottalization that occurs in vowel-initial words that appear in many languages with EGG via other parameters. Garellek[16] presented an

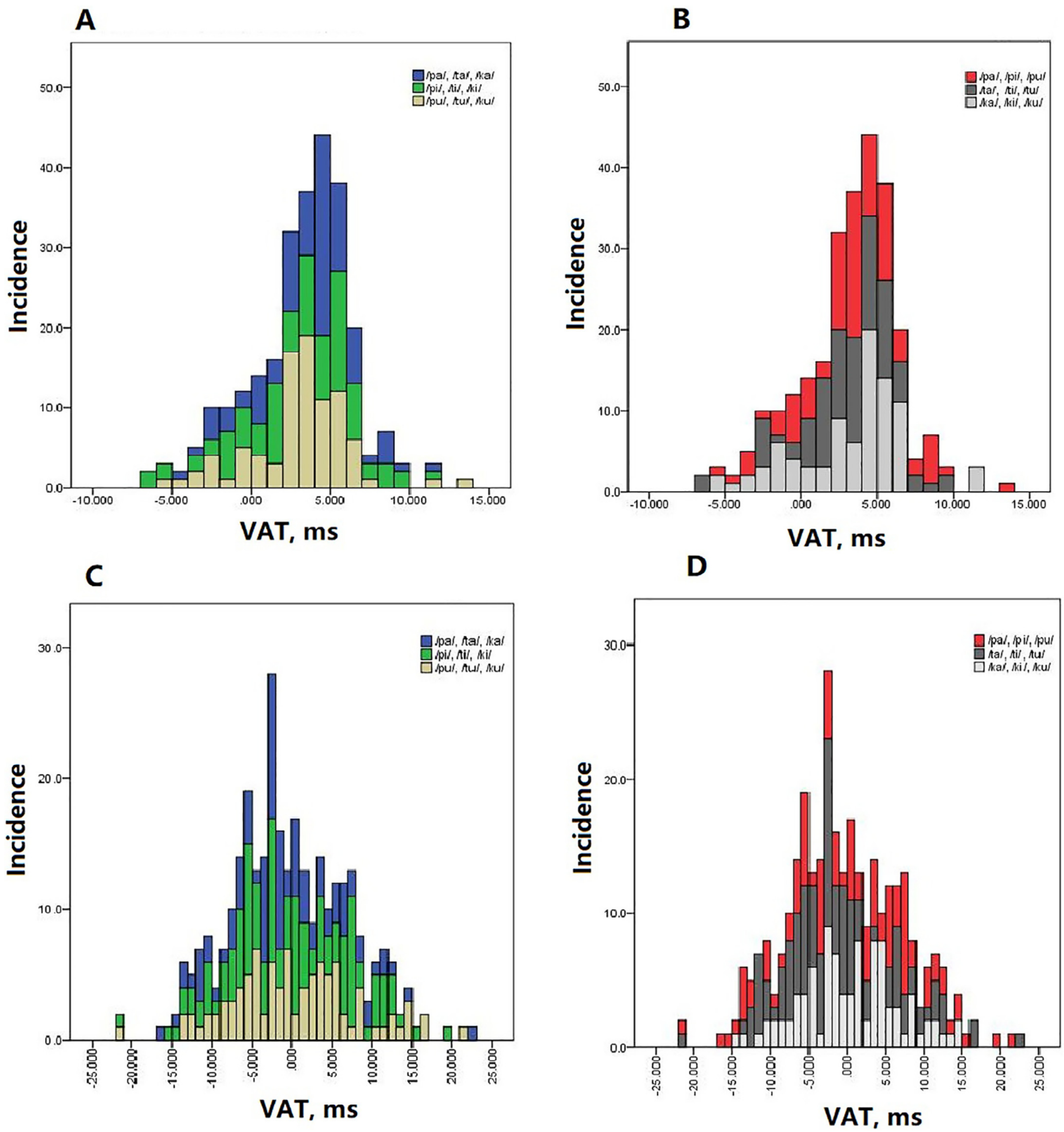

FIGURE 1. Histogram of VAT values as grouped by perceptual analysis and subgroup by monosyllable (bin width = 1 md). **(A)** Same vowel with different consonants in NGS group. **(B)** Same consonant with different vowels in NGS group. **(C)** Same vowel with different consonants in GS group. **(D)** Same consonant with different vowels in GS group.

articulatory (EGG) study to reveal that the prominent glottalization of word-initial vowels, which was realized phonetically as a GS, showed an increase in contact quotient correlated to glottal constriction.

In this study, 635 utterances were not analyzed by age, even the range of age was relatively wide. Up to now, few researches provide the evidence in relation to variation in VAT with age. Certainly, the age issue will be taken into account in future research.

By regarding stimuli, there might well have been different VAT values when single words of Mandarin were embedded in a phrase or sentence. Moreover, there is literature

TABLE 5.
Descriptive Statistics and Independent *t* Test of VAT for All Tokens Categorized by Perceptual Analysis

	VAT					
Group (n)	Mean (SD)	SE of Mean	95% Confidence Interval	Minimum	Maximum	*t*-Value
NGS group (263)	3.19 (3.28)	0.202	2.80–3.59	−6.89	13.15	7.326*
GS group (312)	−0.25 (7.52)	0.425	−1.10 to 0.58	−21.93	22.49	

* $P < 0.001$.
Abbreviations: SD, standard deviation; SE, standard error.

demonstrating the results of the VAT value of single word in phrase (disyllable in Cantonese). However, in this preliminary study about VAT value in patients with cleft palate, only monosyllables were focused on. The next step in this research is to extend the method and analysis to words embedded in phrases and sentences.

As far as phonological processes other than GS are concerned, it is possible that there were other processes represented in the 499 utterances disregarded by the judges, but these also were not the focus of the study.

The comparison of mean VAT values in GS and NGS tokens support the hypothesis that the glottis is constricted during prephonatory period, suggesting that the unaspirated syllable-initial consonants with the negative value of VAT are easier to realize phonetically as GSs in patients with cleft palate in Mandarin. The combination of perceptual and physiological analysis is also required to better explore this issue.

CONCLUSION

The purpose of this study was to investigate how VAT varies when patients with cleft palate with and without GS articulate unaspirated monosyllables in Mandarin. VAT values appear to be sensitively decreased in Mandarin patients with cleft palate with GS compared to those without GS. Our findings confirmed the sensitivity of VAT to measure the existence of GS in patients with cleft palate in Mandarin and suggested its potential application as an objective and quantitative clinical measurement.

Acknowledgments

The authors thank R. J. Baken for the authorization of VAT analysis software to Peking University.

REFERENCES

1. Stengelhofen J. Nature and causes of communication difficulties. In: Stengelhofen J, ed. *Cleft Palate: Nature and Remediation of Communication Problems*. Edinburgh: Churchill Livingstone; 1989:1–30.
2. Prathanee B, Pumnum T, Seepuaham C, et al. Five-year speech and language outcomes in children with cleft lip-palate. *J Craniomaxillofac Surg*. 2016;44:1553–1560.
3. Kummer AW. Anatomy and physiology: the orofacial structures and velopharyngeal valve, cleft palate and craniofacial anomalies, the effects on speech and resonance. *Cleft Palate and Craniofacial Anomalies: The Effects on Speech and Resonance*. San Diego, CA: Singular; 2001.
4. Warren DW. Compensatory speech behaviors in individuals with cleft palate: a regulation/control phenomenon? *Cleft Palate J*. 1986;23:251–260.
5. Sell D, Harding A, Grunwell P. A screening assessment of cleft palate speech (Great Ormond Street Speech Assessment). *Int J Lang Commun Disord*. 1994;29:1.
6. Konst EM, Rietveld T, Peters HF, et al. Use of a perceptual evaluation instrument to assess the effects of infant orthopedics on the speech of toddlers with cleft lip and palate. *Cleft Palate Craniofac J*. 2003;40:597.
7. Kuehn DP, Moller KT. Speech and language issues in the cleft palate population: the state of the art. *Cleft Palate Craniofac J*. 2000;37:348.
8. Whitehill TL. Assessing intelligibility in speakers with cleft palate: a critical review of the literature. *Cleft Palate Craniofac J*. 2002;39:50–58.
9. Kent RD. Hearing and believing: some limits to the auditory-perceptual assessment of speech and voice disorders. *Am J Speech Lang Pathol*. 1996;5:7–23.
10. Howard S. Compensatory articulatory behaviours in adolescents with cleft palate: comparing the perceptual and instrumental evidence. *Clin Linguist Phon*. 2004;18:313–340.
11. Fabre P. Un procédé électrique percutané d'inscription de l'accolement glottique au cours de la phonation: glottographie de haute fréquence. *Bull Acad Nat Méd*. 1957;141:66–69.
12. Taptapova SL, Ermakova J. [Electroglottography studies in patients with congenital cleft palate]. *Rev Laryngol Otol Rhinol (Bord)*. 1979;100:177–181.
13. Hamming KK, Finkelstein M, Sidman JD. Hoarseness in children with cleft palate. *Otolaryngol Head Neck Surg*. 2009;140:902–906.
14. Robison JG, Otteson TD. Prevalence of hoarseness in the cleft palate population. *Arch Otolaryngol Head Neck Surg*. 2011;137:74–77.
15. Zajac DJ, Linville RN. Voice perturbations of children with perceived nasality and hoarseness. *Cleft Palate J*. 1989;26:226.
16. Garellek M. Voice quality strengthening and glottalization. *J Phon*. 2014;45:106–113.
17. Pak M. How allomorphic is English article allomorphy? *Glossa*. 2016;1:20.
18. Solé M-J. Articulatory adjustments in initial voiced stops in Spanish, French and English. *J Phon*. 2018;66:217–241.
19. Shen XR. The acoustic performances of glottal stop. *Stud Lang Linguist*. 2010:35–39.
20. Baken RJ, Orlikoff RF. Estimating vocal fold adduction time from EGG and acoustic records. *Programme and Abstract Book: 24th IALP Congress*. Vol. 15, 1998 Amsterdam.
21. Orlikoff RF, Deliyski DD, Baken RJ, et al. Validation of a glottographic measure of vocal attack. *J Voice*. 2009;23:164–168.
22. Roark RM, Watson BC, Baken RJ. A figure of merit for vocal attack time measurement. *J Voice*. 2012;26:8–11.
23. Watson BC, Baken RJ, Roark RM, et al. Effect of fundamental frequency at voice onset on vocal attack time. *J Voice*. 2013;27:273–277.
24. Watson BC, Baken RJ, Roark RM. Effect of voice onset type on vocal attack time. *J Voice*. 2016;30:11–14.
25. Ma EP, Baken RJ, Roark RM, et al. Effect of tones on vocal attack time in Cantonese speakers. *J Voice*. 2012;26:670.e1.
26. Lee T-YT. The effects of vocal intensity on vocal attack time (VAT). Pokfulam, Hong Kong: The University of Hong Kong. 2010 李芷茵.
27. Tam K-CB. Effects of straw phonation on vocal attack time in Cantonese speakers. Pokfulam, Hong Kong: The University of Hong Kong. 2013 譚家駿.

28. Zhang R, Baken RJ, Kong J. Vocal attack time of different pitch levels and vowels in Mandarin. *J Voice*. 2015;29:542–547.
29. Kong J, Zhang R. VAT of the lexical tones in Mandarin Chinese. *J Chin Linguist*. 2017;45:275–289.
30. Roark RM, Watson BC, Baken RJ. Measures of vocal attack time for healthy young adults. *J Voice*. 2012;26:12–17.
31. Paal S, Reulbach U, Strobel-Schwarthoff K, et al. Evaluation of speech disorders in children with cleft lip and palate. *J Orofac Orthop*. 2005;66:270–278.
32. Dowden PA. *The effects of listener training on the speech intelligibility of severely dysarthric individuals*. 1992;2814.
33. Hikita R, Miyamoto JJ, Ono T, et al. Activation patterns in the auditory association area involved in glottal stop perception. *J Oral Biosci*. 2013;55:34–39.
34. Chen R, Ma L, Zhu H. The phonological characteristics and correction of glottal stop after cleft palate surgery. *Chinese Journal of Stomatology*. 2002;37:191–193.
35. Jiang C, Whitehill TL, McPherson B, et al. Consonant accuracy in Mandarin-speaking children with repaired cleft palate. *Int J Pediatr Otorhinolaryngol*. 2015;79:2270–2276.
36. Moore P. Motion picture studies of the vocal folds and vocal attack. *J Speech Disord*. 1938;3:235–238.
37. Hirose H, Gay T. Laryngeal control in vocal attack. *Folia Phoniatr Logop*. 1973;25:203–213.
38. Cerom JL, Macedo Cde C, Feniman MR. Can peripheral hearing justify the speech disorders in children with operated cleft palate? *Int Arch Otorhinolaryngol*. 2014;18:27–35.