Original Paper

Folia Phoniatrica et Logopaedica

Folia Phoniatr Logop 2012;64:290–296
DOI: 10.1159/000346861

Published online: March 5, 2013

# Design and Evaluation of an Electrolarynx with Tonal Control Function for Mandarin

Congying Wan Erqiang Wang Liang Wu Supin Wang Mingxi Wan

Key Laboratory of Biomedical Information Engineering of Ministry of Education, Xi'an Jiaotong University, Xi'an, PR China

**Key Words**
Electrolarynx · Mandarin · Tonal control

**Abstract**
***Objective:*** To improve the speech quality of laryngectomized speakers of Mandarin, we designed an electrolarynx with tonal control function (tone-EL) by using the movement of a trackball. The performance of the tone-EL in producing Mandarin was evaluated. ***Methods:*** The performance of tone-EL was evaluated via a listening experiment. The perceptual accuracies of monosyllabic words, different categories of phrases with different cues were measured and compared with that produced with a conventional EL. The acceptability of sentences was also measured. ***Results:*** The perceptual accuracies of monosyllabic words and tones associated with the words produced with the tone-EL were significantly higher than those of monotonic EL speech. The perceptual accuracy of phrases was significantly higher than in monotonic EL speech when no categorical cue for listeners was provided, whereas they were at a close level when categorical information of the phrases was provided in advance. The acceptability of sentences was higher than that of monotonic EL speech. ***Conclusion:*** Using the tone-EL can produce Mandarin tones effectively and the quality of rehabilitated Mandarin is better than using a conventional-EL. To balance speech intelligibility and ease of use, a strategy producing tonal speech when few informative cues are available to listeners is preferred.



## Introduction

Laryngectomy, which is an important treatment for laryngeal cancer and trauma, leads to inability of phonation because of the removal of the entire larynx. Since the articulators are still functional, an electrolarynx (EL) can be used to provide a substituted voice source. The EL is an external vibration-generating device, which transmits vibration to the vocal tract through the neck tissue or the oral cavity [1]. Compared with other prostheses, the EL is easier to use and can produce longer sentences without special care [2, 3]. Therefore, the EL is the most commonly used prosthesis for voice rehabilitation [4].

When using an EL, the pitch of EL is preset at a certain level before phonation, but stays steady during voicing. The lack of pitch variation leads to a monotonic sound, which is considered a major flaw of EL speech. Therefore EL devices with real-time pitch adjustment were developed to improve the quality of EL speech. Some com-

KARGER

E-Mail karger@karger.com
www.karger.com/fpl


1021–7762/13/0646–0290$38.00/0

Prof. Mingxi Wan, PhD
Key Laboratory of Biomedical Information Engineering of Ministry of Education
Xi'an Jiaotong University, 28, Xianning West Road
Xi'an, Shaanxi 710049 (PR China)
E-Mail mxwan@mail.xjtu.edu.cn

mercial ELs apply a pressure-sensitive transducer to convert finger pressure to a frequency contour (e.g. TruTone EL, Griffin Laboratories). Takahashi et al. [5, 6] used a miniature fingertip switch to control the accent of a denture-based EL according to the amount of finger pressure. Such finger pressure ELs are characterized by simplicity and availability, allowing for rapid pitch changes. Uemi et al. [7] placed an air pressure sensor on the neck stoma of laryngectomees to obtain pitch modulation from expiration pressure. The expiration-control EL produced more natural intonation for esophageal or tracheoesophageal speakers after training. Goldstein et al. [8] designed a hands-free EL triggered by neck strap muscle electromyographic (EMG) activity. An envelope obtained from the rectified EMG signal by applying a 1-Hz low-pass filter was used to modulate pitch. After training, users could produce intonation distinctions in their speech [9].

Although these improved EL devices proved to be effective in generating intonation or stress, they have not been applied in speech rehabilitation of tonal languages, in which tone plays an important role by signifying meaning. In Mandarin, each monosyllable is produced with a distinctive tone, which makes use of fundamental frequency ($F_0$) to distinguish minimal word pairs that are not differentiated by segmental cues [10]. For example, the monosyllable *ma* with the four types of Mandarin tones high-level (HL), middle-rising (MR), falling-rising (FR), and high-falling (HF) means 'mother', 'numb', 'horse', and 'curse', respectively. Moreover, the pitch contour is considered to be the primary contribution to the perception of meaning in tonal languages [11]. Therefore, providing tonal control will improve the intelligibility of EL Mandarin speech a lot. However, the performance of producing tones with current EL devices is doubtful. For finger pressure ELs, the pitch changes during speaking are difficult to match with finger pressure control [4]. The expiration EL is hard for proficient EL users to master, because they are used to producing voice in the absence of airflow [12]. For EMG-EL, the signal containing energy below 1 Hz cannot modulate frequency fast enough to produce tones.

Aiming at providing the four Mandarin tones, we designed an EL device with tonal control function (tone-EL). The movement of a trackball was used to control the frequency of a wearable EL vibrator. The trackball provides a simple and general interface for users, who can produce the four Mandarin tones by manipulating the trackball just like drawing the pitch contours during voicing. In this paper, the design and implementation of the device are introduced first. Then we assess the perceptual characteristics of speech produced with tone-EL, including the perceptual accuracy of monosyllabic words, the perceptual accuracy of Mandarin phrases under different conditions, and the acceptability of sentences. Through these tests, we discuss the principle of generating a strategy that could balance the speech quality and the ease of use under different conditions.

## Methods

*Design and Implementation of the Tone-EL*

The tone-EL consists of three parts: a trackball as a manual control unit; a control circuit for converting the movement of the trackball to frequency adjustment and output driving signals, and a wearable EL vibrator. The trackball mouse (Dongli-E; Shenzhen, China) is used to provide manual control. The mouse is suitable for a user to hold and manipulate the trackball with the thumb. The movement of the trackball is converted to digital data and then transferred to a system on a programmable chip, which is built on an FPGA board (Cyclone II 2C35; Altera, Calif., USA). By decoding the digital data, the SOPC obtains the movements of the trackball on x (horizontal) and y (longitude) axes. The EL vibrator is turned on when the trackball is active. In the meantime, the movement on the y axis is used to adjust the frequency of a direct digital synthesizer. After power amplification, the signal drives the wearable vibrator (Xinyu; Daqing, China) to produce a substitute voice source.

When using the tone-EL, the user puts on the wearable vibrator, holds the mouse, and manipulates the trackball to switch the EL and adjust the pitch contour. When coordinated with articulation, tonal speech is produced. Besides the tonal control mode, the device also provides a monotonic mode for users, who can press a button to produce monotonic EL speech like a conventional EL.

*Frequency Control*

The four Mandarin tones, HL, MR, FR, and HF, are characterized by fluctuation of $F_0$ as a function of time. In the implementation of the tone-EL, $F_0$ variation is determined by the direction of trackball movement along the y axis. To produce specific tones, the user manipulated the trackball for drawing the pitch contours. The $F_0$ was set to 110 Hz at the initiation of voicing, and was adjusted in steps of 1 Hz per 5 ms, equaling 200 Hz/s. These settings were based on informal tests during preliminary sessions. A slower rate of $F_0$ adjustment resulted in inefficient $F_0$ height variation. A faster rate easily caused the $F_0$ to exceed the range of 40–200 Hz during one voicing. When $F_0$ was below 40 Hz, the voice produced with the device sounded like discrete impulses. When $F_0$ was over 200 Hz, the voice was not loud enough. The adjusting interval of 5 ms was sufficient for producing smooth $F_0$ contours without too much requirement of ROM in the direct digital synthesizer. These settings were deemed effective in producing distinct tones. After one monosyllabic word was finished, the user stopped manipulating the mouse and waited for 100 ms to reset the $F_0$. Figure 1 shows a comparison of $F_0$ contours for the syllable /ma/ produced with normal voice, conventional EL, and tone-EL, respectively.

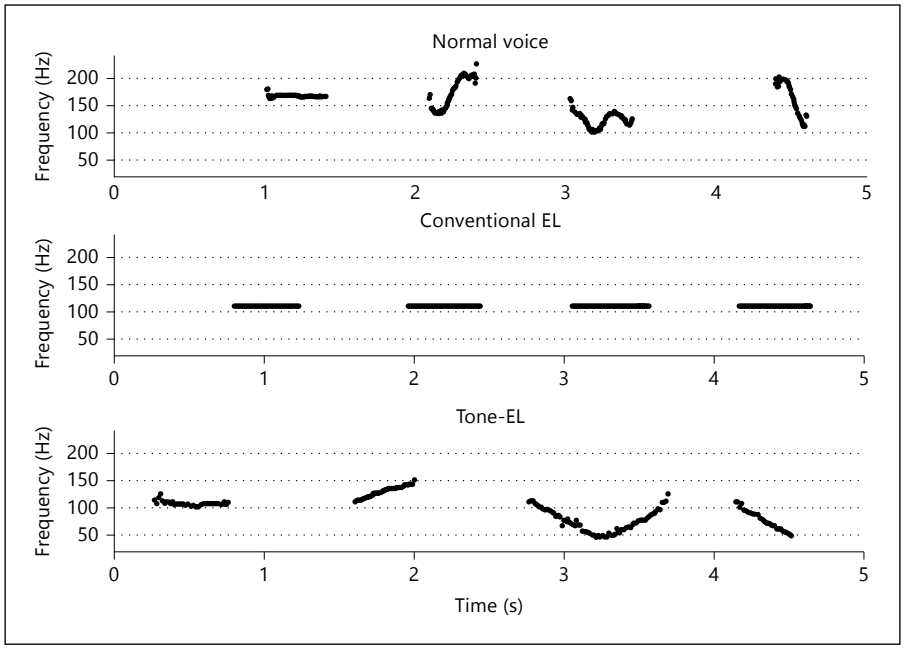


**Fig. 1.** $F_0$ contours of the syllable /ma/ produced with normal voice, a conventional EL, and the tone-EL.

*Participants: Speakers*

Two persons with laryngectomy were invited to participate in the study to produce voice with the tone-EL. The persons were male, 76 and 68 years old; both had their larynx removed 4 years prior to the study. They had used conventional EL for communication for more than 3 years. They had no reported history of speech or language problems except that associated with laryngectomy. Both participants are native Mandarin speakers, who are literate and had no difficulty reading the speech material used in the study.

The persons were introduced to the use of the tone-EL first, and then were asked to practice producing the four tones with the tone-EL. In addition to the audio feedback, a virtual instrument programed with LabVIEW (NI Instruments, Austin, Tex., USA) was used to provide visual feedback. The virtual instrument acquired the audio signal, calculated the $F_0$ of the voice produced by users, and plotted the $F_0$ contours on the screen. After practicing for about 20 min, both users were familiar with the use of the tone-EL and could produce the four tones with high quality.

*Participants: Listeners*

Eight native speakers of Mandarin Chinese participated in the perceptual experiment. Their ages varied from 23 to 29 years, with a mean age of 25.8 years. All the listeners had at least a high school education, and they were able to correctly comprehend the speech material used in the experiment. The listeners had no prior experience of listening to EL speech. Each listener declared that he or she had no known history of hearing problems.

*Reading Materials*

Four sets of reading materials were prepared for the 2 speakers to read. Each set of materials consisted of 75 monosyllabic words, 60 phrases, and 10 sentences. The monosyllabic words were derived from a list randomly chosen from the China National Standards of the acoustic-speech articulation testing method (GB/T 15508-1995), which contains ten equivalent lists of monosyllabic words. A phrase was a combination of two or more monosyllabic words. Each set contained three categories of phrases, including 20 numbers, 20 adjectives, and 20 proper nouns. The numbers were integers randomly chosen from 1–999. The adjectives are frequently used in daily life. The proper nouns were names of well-known persons or places. The average length of phrases were 2.6, 2.1, 3.4 monosyllabic words for numbers, adjectives and proper nouns, respectively. Ten sentences were chosen from the newspaper and the average length was 9.6 monosyllabic words. Each speaker was asked to read two of the sets of materials with a conventional EL and a tone-EL, respectively. The audio signal was recorded with a high-quality microphone, and each sound was saved as an mp3 file with a sampling rate of 44,100 per second.

*Listening Task*

The listening test was done in a soundproof room. The listeners were asked to complete four tasks. Four sets of stimuli corresponding to the four tasks were constructed based on the stimuli produced by the speakers. Within each set, all the sounds, whether produced with conventional EL or tone-EL, were mixed together and randomly sorted. The audio files were presented independently to listeners via binaural earphones at a comfortable volume. Listeners were allowed to listen to the sounds as many times as needed.

The first set of stimuli consisted of all the monosyllabic words ($75 \times 4 = 300$ words). In this task, the listeners were asked to write down the words they heard. Two kinds of perceptual accuracies were calculated. The accuracy of segment perception was the percentage of correctly identified segmental features regardless of tone, and the accuracy of word perception was the percentage of both correctly identified segmental features and tones. Confusion matrices of responses to Mandarin tones associated with the monosyllabic words were also constructed to indicate the pattern of perception for different tones.

The second and third tasks were designed to assess the perceptual accuracy of phrases produced with the tone-EL under different conditions. The phrases (60 × 4 = 240 phrases) were divided into two equal parts to form the second and third sets of stimuli. Thus each set consisted of 120 phrases, half of which were produced with conventional EL (60 in total, including 20 numbers, 20 adjectives, and 20 proper nouns) and the rest were produced with tone-EL. The second set was played to listeners without any information in advance. For the third set, the listeners were given categorical information about the phrase (number, adjective, name of a person, or name of a place) before listening to each phrase. The different steps were adopted to simulate the comprehension of EL speech in practical conversation with and without hints from previous contexts. The listeners were asked to write down the phrases they heard. The accuracy of phrase perception was calculated as the percentage of correct responses (all monosyllabic words that were correctly identified within one phrase).

The fourth task was to assess the acceptability of tone-EL speech. The fourth set of stimuli consisted of all the sentences (10 × 4 = 40 sentences). Listeners were asked to score the sentences according to the naturalness of speech. A mean opinion score, which was a five-point scale (1 = bad; 2 = poor; 3 = common; 4 = good; 5 = excellent), was used. The overall acceptability was the average score achieved across all the speech materials. Prior to beginning the task, the listeners listened to all the 40 sentences to roughly rank them.

## Results

Figure 2 shows the perceptual accuracies of segments and monosyllabic words. There was no significant difference in segment perceptual accuracy between conventional EL speech and tone-EL speech (t tests, $p = 0.398$). However, the word perceptual accuracy of tone-EL speech was significantly better than that of conventional EL speech (one-sided t tests, $p < 0.001$).

Table 1 shows the confusion matrices for tones associated with monosyllabic words produced with conventional EL and tone-EL. Post hoc test showed that the perceptual accuracy of HL associated with monosyllabic words produced with conventional EL was significantly better than that of the other tones ($p < 0.001$). For tones associated with monosyllabic words produced with the tone-EL, the perceptual accuracy of HL and HF did not show a significant difference ($p = 0.359$), the perceptual accuracy of FR was not significantly different from that of MR ($p = 0.095$), but was significantly greater than that of HL ($p = 0.003$) and HF ($p = 0.025$).

Table 2 shows the perceptual accuracy of different types of phrases produced with conventional EL and tone-EL. When no prior information was provided for the listeners, the interaction between tonal information and categorical information was not significant [two-way ANOVA, $F(2, 48) = 1.493$, $p = 0.236$], tonal information significantly affected perceptual accuracy [two-way ANOVA, $F(1, 48) = 31.176$, $p < 0.001$]. When listeners were provided with categorical information in advance, the interaction between tonal information and categorical information was not significant [two-way ANOVA, $F(2, 48) = 2.007$, $p = 0.147$], the influence that tonal information had on the perceptual accuracy was not significant either [two-way ANOVA, $F(1, 48) = 2.404$, $p = 0.129$].

The acceptability of conventional EL speech was 2.1 (SD = 0.47), while the acceptability of tone-EL speech was 3.8 (SD = 0.54). The acceptability of tone-EL speech was significantly larger than that of conventional EL speech (one-sided t tests, $p < 0.001$).

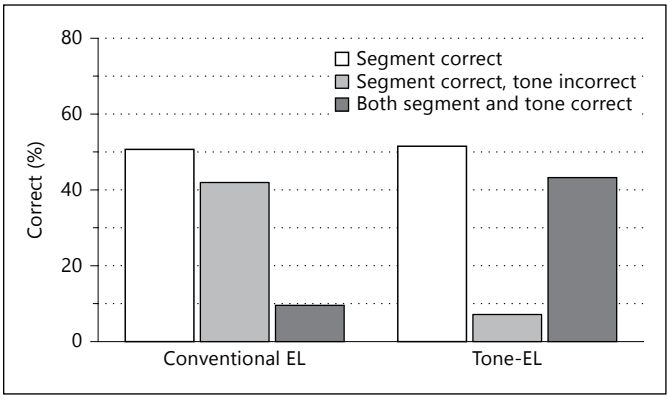


**Fig. 2.** Perceptual accuracies of monosyllabic words.

**Table 1.** Confusion matrix of responses to Mandarin tones associated with single words

| Stimulus tones | Response tones | | | |
|---|---|---|---|---|
| | HL | MR | FR | HF |
| Produced with conventional EL | | | | |
| HL | *92.8 (5.4)* | 8.9 (2.5) | 0 (0) | 6.3 (5.9) |
| MR | 73.6 (19.4) | *13.9 (11.1)* | 1.4 (2.6) | 11.1 (11.9) |
| FR | 83.3 (16.3) | 9.2 (8.7) | *1.7 (3.1)* | 5.8 (9.0) |
| HF | 74.6 (18.9) | 6.4 (6.0) | 5.7 (5.3) | *13.3 (14.3)* |
| Produced with the tone-EL | | | | |
| HL | *80.8 (4.3)* | 8.3 (4.7) | 7.5 (6.6) | 3.4 (5.0) |
| MR | 6.6 (6.6) | *86.8 (10.8)* | 2.2 (4.4) | 4.4 (4.2) |
| FR | 3.3 (5.0) | 1.7 (3.1) | *93.3 (6.2)* | 1.7 (3.1) |
| HF | 2.2 (3.3) | 0.9 (1.6) | 12.5 (6.3) | *84.4 (7.6)* |

Standard deviations in parentheses.



Folia Phoniatr Logop 2012;64:290–296
DOI: 10.1159/000346861

**Table 2.** Perceptual accuracy of different types of phrases produced with conventional EL and the tone-EL

| | No prior information provided | | | Category information of phrases provided | | |
|---|---|---|---|---|---|---|
| | adjectives | numbers | proper nouns | adjectives | numbers | proper nouns |
| Conventional EL | 24.4 (6.2) | 63.8 (9.5) | 39.4 (10.2) | 48.8 (7.9) | 87.7 (9.2) | 77.8 (7.5) |
| Tone-EL | 35.0 (8.5) | 84.4 (7.8) | 51.2 (10.6) | 58.7 (9.5) | 87.5 (4.6) | 79.4 (7.6) |

Standard deviations in parentheses.

## Discussion

*Design of the Tone-EL*

In this study, aiming at producing tones for speech rehabilitation of Mandarin, we designed and implemented an EL with tonal control function. A trackball mouse was used as the interface for $F_0$ control. The users can adjust the $F_0$ contours by controlling the movement of the trackball during voicing. To provide easy use, settings including fixed initial $F_0$ and fixed rate of $F_0$ variation were adopted. The advantage of fixed settings is the simplicity and effectiveness of producing distinct tones. The users just need to focus on the direction of trackball movement and the duration of manipulation during phonation with the tone-EL. However, the settings result in several limitations of the tone-EL. First, the contrasts between different tones associated with words produced with the tone-EL are unlike those of tones associated with normal voice. The influence that the settings have on speech perception is discussed later. Second, the user cannot produce continuous speech, because he or she has to wait for at least 100 ms after finishing each monosyllabic word to reset the initial $F_0$ for the next word. This may affect the naturalness of connected speech. Third, the device is not suitable for producing intonation and stress due to the relatively fast rate of $F_0$ adjustment and the fixed settings.

The fixed settings are a compromise between speech quality, easy use of the interface, and the implementation of the device. The break between monosyllabic words seems to be unavoidable in the implementation of an EL device with Mandarin tone control function, because the $F_0$ contours of sequential words may be discontinuous. An adjusting rate proportional to the velocity of the trackball may produce $F_0$ contours more approximated to the normal voice. In addition, the proportional $F_0$-adjusting rate can provide more flexible $F_0$ patterns so that it may enable the user to produce additional intonation or stress on sentences. However, users of such an interface need to make a special effort to precisely control the trackball and may find it difficult to master. Moreover, how can an ordinary user without knowledge of phonology specify the $F_0$ contours combining tone patterns and intonation or stress? The fixed initial $F_0$ is mainly due to the trackball which cannot provide the information of its initial position. This problem may be solved by using interfaces sensitive to initial conditions.

*Speech Perception*

Due to the lack of pitch adjustment, most tones associated with monosyllabic words produced with a conventional EL were identified as HL. The tones associated with monosyllabic words produced with a tone-EL were correctly identified with high accuracy. This result indicates that the users could produce distinct tones with the tone-EL effectively. FR was identified with the highest accuracy. This may due to the distinctive pitch contour of FR, which contains a turning point. Accordingly, the lack of the turning point leads to the least perceptual accuracy of FR produced with a conventional EL. HL was identified with the least accuracy. This may be mainly due to the manipulation of the trackball. The $F_0$ adjustment is sensitive to the movement of the trackball to satisfy the requirement of fast $F_0$ change. It may be difficult for the user to produce a flat pitch contour without fluctuations. This can be improved by using the button instead of the trackball to produce HL. The most frequent confusion occurred between HF and FR. This may be related to the fixed settings. HF produced with a normal voice starts with a higher level of frequency than FR, and falls to a low level rapidly. However, HF produced with tone-EL starts with the same frequency as FR, and falls to the low level with the same rate of frequency variation as the falling part of FR. As a result, the $F_0$ contour of HF is very close to the falling part of the $F_0$ contour of FR. Therefore, the HF produced with the tone-EL is misidentified as FR more often than the rest of the tones.

The results of intelligibility of monosyllabic words produced with a conventional EL (15.2%) and with the

tone-EL (49.8%) are close to the results reported by Zhang et al. [13]. In their research, the intelligibility of monosyllabic words excited by triangular wave at fixed $F_0$ was 16.3%, while the intelligibility of words excited by triangular wave with normal tones was 50.3%. Although the tones are not helpful for the perception of syllables, high accuracy in tone perception leads to a much higher perceptual accuracy for monosyllabic words.

When there was no hint for listeners, the listeners identified phrases produced with the tone-EL more accurately than phrases produced with a conventional EL. The perceptual accuracy for phrases of different types was improved 11.4–20.6% by using the tone-EL instead of a conventional EL. The improvement was mainly due to the additional cues provided by tones. However, the improvement was much less than that reported by Zhang et al. [13], where the intelligibility of phrases improved about 40% by providing natural tones. This may be explained as follows. The contrast between the four tones produced with the tone-EL is different from that of the normal voice. It has been found that tonal identification does not only depend on intrinsic acoustic information of frequency variation with time, but also on contrast with other tones in the utterance [14]. For tone-EL speech, the four tones start with the same $F_0$, and the $F_0$ contours vary at a fixed rate. Therefore, the perception of tones of phrases is not as accurate as that of monosyllabic words. In addition, the lack of tone sandhi which occurs frequently within phrases might affect the perception of phrases.

The result of an acceptability test indicates that speech produced with the tone-EL is more acceptable than conventional EL speech. Although the tone-EL produces sentences consisting of discrete words, listeners prefer discontinuous tonal speech rather than continuous monotonic speech.

*Strategy of Using the Tone-EL*

The intelligibility of tone-EL speech is higher than that of conventional EL speech under most conditions or no less than that of conventional EL under some conditions. Moreover, the acceptability of tone-EL speech is significantly higher than that of conventional EL speech. Perceptual results indicate that using tone-EL can effectively improve the quality of EL speech. However, the use of the tone-EL requires more concentration than the use of a conventional EL. Both of the 2 laryngectomized users invited in this study claimed that the main effort was on predetermining the lexical tones associated with words. This may imply that there is an optimal strategy for using the tone-EL to achieve a balance between speech quality and ease of use in communication. Therefore, the device is designed to provide both tonal mode and monotonic mode for users to switch flexibly.

To achieve the best speech quality without considering the effort on the use of the tone-EL, always producing the tone for each word is the optimal choice. However, it is not necessary to produce every tone to achieve the best intelligibility in practical communication, because monotonic EL speech can be as intelligible as tonal EL speech with sufficient informative cues. In this study, categorical information was provided as a hint obtained in advance. The informative cues provided by categorical information are of different levels. In Mandarin, a few basic words can express all the numbers, and proper nouns are usually unique phrases with fixed words, while adjectives in Mandarin are more flexible. Therefore, the categorical information might provide enough cues for listeners to identify a number or proper noun, but might not be enough to identify an adjective. When insufficient cues are provided, tonal information remarkably improves the identification of phrases. However, when sufficient cues are provided, tonal information may be dispensable without affecting comprehension. Therefore, the strategy must be based on the level of informative cues that listeners obtain extrinsic and intrinsic to the phrases.

However, we have to admit that the experiment adopted in this study is much simpler than practical communication, in which more types of phrases are used and the hints are provided with various approaches. Therefore, we cannot give a concrete strategy for using the tone-EL, but just point out that the optimal strategy is related to the level of informative cues obtained extrinsic and intrinsic to the phrases. More systematic and concrete research on alaryngeal speech in combination with linguistics will be helpful for guiding the use of the tonal EL. In addition, many other factors, such as different levels of speech quality and the individual proficiency of using the tone-EL, should be taken into consideration in practical use of the tone-EL.

## Conclusion

In this study, we designed and implemented an EL device with tonal control function for Mandarin voice rehabilitation. The movement of a trackball was used to control the vibration $F_0$ of the EL. The user could produce Mandarin tones effectively by manipulating the trackball. The perceptual experiment showed that the tones associated with monosyllabic words produced with the tone-EL were identified with high accuracy; the intelligibility and acceptability of tone-EL speech were significantly im-

proved. However, the improvement that tonal information induced on intelligibility was not significant when listeners obtained sufficient cues in advance. In addition, producing tones manually required more concentration than using a conventional EL. To balance speech quality and ease of use, users were suggested to use the tonal mode or monotonic mode of the tone-EL according to the level of informative cues that the listener could obtain extrinsic and intrinsic to the phrases.


## Acknowledgments

We deeply appreciate Prof. Gary Weismer for his useful suggestions and careful editing of the text. This work was supported in part by the National Natural Science Foundation of China under grant 61271087 and grant 11274250.


## References


1 Barney HL, Haworth FE, Dunn HK: An experimental transistorized artificial larynx. Bell Syst Tech J 1969;38:1337–1356.

2 Lauder E: The laryngectomee and the artificial larynx – a second look. J Speech Hear Disord 1970;35:62–65.

3 Rothman HB: Acoustic analysis of artificial electronic larynx speech; in Seikey A (ed): Electroacoustics Analysis and Enhancement of Alaryngeal Speech. Springfield, Thomas, 1982, pp 95–118.

4 Liu H, Ng ML: Electrolarynx in voice rehabilitation. Auris Nasus Larynx 2007;34:327–332.

5 Takahashi H, Nakao M, Okuas T, Hatamura Y, Kikuchi Y, Kaga K: A voice-generation system using an intra-mouth vibrator. J Artif Organs 2001;4:288–294.

6 Takahashi H, Nakao M, Kikuchi Y, Kaga K: Alaryngeal speech aid using an intra-oral electrolarynx and a miniature fingertip switch. Auris Nasus Larynx 2005;32:157–162.

7 Uemi N, Ifukube T, Takahashi M, Matsushima J: Design of a new electrolarynx having a pitch control function. Proc IEEE Int Workshop on Robot and Human Commun 1994, pp 198–203.

8 Goldstein EA, Heaton JT, Kobler JB, Stanley GB, Hillman RE: Design and implementation of a hands-free electrolarynx device controlled by neck strap muscle electromyographic activity. IEEE Trans Biomed Eng 2004;51:325–332.

9 Goldstein EA, Heaton JT, Stepp CE, Hillman RE: Training effects on speech production using a hands-free electromyographically controlled electrolarynx. J Speech Hear Res 2007;50:335–351.

10 Yiu EM, Hasselt CA, Williams SR, Woo JKS: Speech intelligibility in tone language (Chinese) laryngectomy speakers. Eur J Disord Commun 1994;29:339–347.

11 Liu H, Wan M, Ng ML, Wang S, Lu C: Tonal perceptions in normal laryngeal, esophageal, and electrolaryngeal speech of Mandarin. Folia Phoniatr Logop 2006;58:340–352.

12 Liu H, Wan M, Wang S, Niu H: Aerodynamic characteristics of laryngectomees breathing quietly and speaking with the Electrolarynx. J Voice 2004;18:567–577.

13 Zhang J, Qi S, Song M, Liu Q: On the important role of Chinese tones in speech intelligibility. Chin J Acoust 1981;3:278–283.

14 Moore CB, Jongman A: Speaker normalization in the perception of Mandarin Chinese tones. J Acoust Soc Am 1997;102:1864–1877.