Speech Intelligibility, Acceptability, and Communication-Related Quality of Life in Chinese Alaryngeal Speakers

Ida K.-Y. Law, MPhil; Estella P.-M. Ma, PhD; Edwin M.-L. Yiu, PhD

Objective: To investigate (1) speech intelligibility and acceptability in using 4 different alaryngeal speech methods: esophageal (ES), electrolaryngeal (EL), pneumatic device (PD), and tracheosophageal (TE) speech; and (2) communication-related quality of life (QOL) in the alaryngeal speakers who used these 4 alaryngeal speech methods.

Design: Survey.

Participants: Alaryngeal speakers who had undergone speech rehabilitation and were recruited from the New Voice Club of Hong Kong.

Main Outcome Measures: Speech samples collected from 49 alaryngeal speakers were rated by 6 judges for speech intelligibility and acceptability. The speakers also completed a communication-related QOL questionnaire called the Communication Activity and Participation After Laryngectomy.

Results: We found that the ES and EL speakers showed considerably poorer speech intelligibility and communication-related QOL. The PD speakers demonstrated notably better speech intelligibility and acceptability ratings. However, high intelligibility and acceptability do not necessarily mean better QOL. The TE speakers, who demonstrated only the second highest speech intelligibility and acceptability, showed the best functional QOL.

Conclusion: In speech rehabilitation after laryngectomy, QOL and speech intelligibility and acceptability should be considered together to find a balance that is acceptable to the patient.

levels. Because pitch variations are essential to convey different lexical meanings of Cantonese words, such difficulties would affect speech intelligibility. A number of studies of the Cantonese-speaking alaryngeal population focused on the speech intelligibility of different alaryngeal speech, with particular attention to lexical tone production. In the study by Yiu et al., ES speakers demonstrated the highest speech intelligibility, but TE and PD speakers were more efficient in conveying lexical tones. Similarly, Ching et al. also found PD speakers to be able to produce better pitch variations, and hence better lexical tone production, than those using other types of alaryngeal speech methods, although there was no notable difference in the overall speech intelligibility among the 4 different alaryngeal speech types.

Speech impairments following laryngectomy, including lower speech intelligibility and poorer lexical tone production, can have an impact on the alaryngeal speaker’s communication abilities. The extent to which speech intelligibility in different speech methods has an impact on quality of life (QOL) is largely unknown. Clements et al. investigated ES, EL, and TE speakers’ satisfaction with their speech quality, communication abilities, interaction with others, and general QOL. They found TE speakers to be notably more satisfied with their speech method and to have better speech quality than ES and EL speakers. Nevertheless, no important difference was found in the level of satisfaction with QOL across the 3 alaryngeal speech groups. In general, their participants reported low self-perceived ratings on their abilities to communicate over the telephone, high levels of limitation in communicating with others, and low levels of satisfaction with QOL. However, the small number of communication activities investigated in their study would not be adequate to assess the communication of the alaryngeal speaker population. Moreover, no PD speakers were included in the study because PD is not a commonly used speaking method in Western countries.

In Hong Kong, most patients who have undergone laryngectomy are seen in public hospitals. The TE option is usually discussed with the patients before the surgery. Individuals who are not suitable for the TE option are all given the other 3 alaryngeal speech options. Most patients voluntarily choose the EL speech as their immediate means for communication after surgery. The other 2 options, ES and PD, are also tried by patients during rehabilitation follow-up sessions.

The present study compared the speech intelligibility and acceptability among 4 currently available alaryngeal speech modes, namely, EL, ES, TE, and PD speech using 2 distinct age groups of listeners (25-33 years vs 60-65 years). The study also compared the communication-related QOL among speakers of these 4 alaryngeal speech modes. The information on the speech intelligibility, acceptability, and QOL in different Chinese alaryngeal speech methods would provide health care workers and alaryngeal speakers an opportunity to make informed decisions as to the selection of a more suitable alaryngeal speech method. Although the PD method is not popular in Western societies, the information provided herein will serve to show that it can be a useful option for alaryngeal speakers whose mother tongue is one of those tonal languages.

### METHODS

#### PARTICIPANTS

Fifty-six native Cantonese-speaking alaryngeal speakers recruited from the New Voice Club of Hong Kong participated in the present study. All these participants had stopped attending the regular speech therapy clinics provided by the hospitals, and therefore it can be assumed that they all had achieved their maximum ability in acquiring the new speaking methods. Each participant was given a hearing screening test at octave frequencies ranging from 250 to 8 kHz at a 20-dB hearing level (HL). Seven participants were excluded from the study owing to other medication complications or having failed the hearing screening test. Therefore, only data from 49 alaryngeal speakers, ranging in age from 40 to 77 years (mean[SD] age, 65.37 [8.77] years) and with a postoperative period ranging from 6 to 194 months (mean[SD] duration, 79.22 [38.76] months) were included in the final analysis. Table 1 shows the sex distribution and postoperation time across the 4 alaryngeal speech method groups. All of the EL speakers used the neck-type EL devices, and all of the TE speakers were fitted with the Blom-Singer (InHealth Technologies, Carpinteria, California) valve using digital occlusion. The seemingly low proportion of female alaryngeal speakers indeed reflected the distribution of sex in the alaryngeal-speaking population in Hong Kong. Kruskal-Wallis 1-way analysis of variance (ANOVA) revealed no significant differences in the length of postoperative period (P = .27), age (P = .12), and education level (P = .20) among the 4 alaryngeal speech method groups.

#### PROCEDURES

**Perceptual Evaluations of Speech Intelligibility and Acceptability**

**Speech Material.** The Cantonese Sentence Intelligibility Test (CSIT) was used to assess the speech intelligibility of the par-

---

**Table 1. Sex Distribution and Postoperation Time**

<table>
<thead>
<tr>
<th>Alaryngeal Speech Mode</th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
<th>Mean (SD) Duration</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Esophageal</td>
<td>0</td>
<td>7</td>
<td>7</td>
<td>120.43 (51.23)</td>
<td>24-194</td>
</tr>
<tr>
<td>Electrolarynx</td>
<td>0</td>
<td>14</td>
<td>14</td>
<td>62.79 (62.41)</td>
<td>6-190</td>
</tr>
<tr>
<td>Tracheoesophageal</td>
<td>1</td>
<td>12</td>
<td>13</td>
<td>70.92 (51.23)</td>
<td>20-181</td>
</tr>
<tr>
<td>Pneumatic device</td>
<td>3</td>
<td>12</td>
<td>15</td>
<td>82.53 (60.44)</td>
<td>7-192</td>
</tr>
</tbody>
</table>

---

©2009 American Medical Association. All rights reserved.
participants at sentence level. The CSIT was developed from the Assessment of Intelligibility of Dysarthric Speech (AIDS). In the CSIT, there are a total of 1100 daily Chinese (Cantonese) sentences (100 sentences, each with 5-15 words, for each sentence length set), which make up the master pool. These sentences do not contain quotations, reduplication, parentheses, proper names, or numbers larger than 10. For each speaker, a unique set of 22 sentences was randomly selected within the master pool, with a sentence length that varied from 5 to 15 words. Two sentences were selected for each sentence length, and they were all printed out on a single sheet of A4 paper (21 × 29 cm) in a 22-point font size.

Recording Procedures. Recording of speech samples took place in a quiet room located at the New Voice Club of Hong Kong with the background noise level kept below 45 dB (A-weighted) (hereinafter noise level) as measured by a sound level meter. All the speech samples were recorded using a Sony MZ-R909 Mini-Disc Digital recorder (Tokyo, Japan) with a Shure SM48 dynamic microphone (Niles, Illinois) held 10 cm away from the center of the speaker’s mouth. Each speaker was given a brief practice period to become familiar with the speech materials, recording instructions, and procedures before the actual recording. To ensure that the quality of the recorded speech samples was representative of the speaker’s daily communication speech, a randomly selected sample from the recordings of each participant was played back to that individual. They were asked to judge if the recorded speech sample was representative of the quality of their daily communication speech. If the played-back sample was judged to be not representative by the individual speaker, the recording procedure was repeated until a representative sample was recorded.

In the recording of the sentences from the CSIT, the sentence list to be read was placed in front of each individual. The participants were instructed to read each of the 22 sentences as clearly as possible. The order of sentences to be recorded was randomized to avoid any possible order effect. Also, to avoid any reading errors owing to misreading of words, the following procedures were used. First, the investigator (I.K.-Y.L.) would indicate to the speaker each sentence to be read by pointing to the respective sentence. The speaker was asked to read the indicated sentence together with the investigator once. Then the investigator read the sentence aloud alone once, followed by the speaker, who read the sentence aloud alone once. This final attempt by the speaker was recorded. If any obvious reading error was noted in the recording, that sentence would be recorded again.

Listeners. Six native Cantonese-speaking adults served as judges to transcribe and rate the speech samples. Three of them were 25 to 33 years old (younger judge group), and 3 others were 60 to 65 years old (older judge group). They were asked to transcribe and provide perceptual ratings of speech impairment for all the speech samples. All judges had no prior exposure to individuals who had undergone a laryngectomy and were alaryngeal speakers. They were not aware of the content of the speech materials used in the study before the listening tasks. All of them passed a hearing screening test at a threshold level of 20-dB HL at each octave frequency (250-8000 Hz).

Listening Tasks

A Sony CMT-JST MD high-fidelity sound system was used to play all the recorded speech samples to each individual judge with the speakers at a distance of 2 m in a quiet room (with a background noise level <45 dBA as measured by a sound level meter). The intensity level of the playback was adjusted by individual judges to a comfortable level. The order of the alaryngeal speakers and the recorded sentences of each speaker was played back at a randomized order across listeners to avoid any possible order effect.

Each of the 6 judges scored each alaryngeal speaker individually. They were asked to judge speech samples of the CSIT first. They were instructed to orthographically transcribe the 22 sentences and rate the severity of speech impairment of each speaker based on the speech samples on an 11-point (0-10) equal-appearing interval (EAI) scale with 0 corresponding to no speech impairment and 10 corresponding to very severe speech impairment. In the orthographic transcription task, the judges were first asked to listen to all 22 sentences together once and then listen to each of the sentences a second time, pausing between 2 sentences whenever necessary, to transcribe each sentence word by word on the score sheets. The guideline suggested by Yorkston and Beukelman was followed; that is, judges were not allowed to listen to each sentence more than 2 times and were encouraged to guess at words that were not completely understood.

Each judge was required to repeat these tasks for 40 randomly selected speaker sample sets (ie, 80% of the total number of speaker sample sets) 2 weeks after their first listening task. This procedure was to establish the intrajudge reliability. The interjudge reliability was established by correlating the intelligibility scores transcribed by an individual judge with the other 2 judges within the younger and older judge groups.

Communication-Related QOL

Each participant also completed the Communication Activity and Participation After Laryngectomy (CAPAL) questionnaire (Figure) to evaluate the extent of communication-related QOL deterioration in alaryngeal speakers. This profile was developed using the International Classification of Functioning, Disability, and Health (ICF) framework, which evaluates the communication-related QOL in individuals who have undergone laryngectomy and experienced 3 levels of disablement, namely, speech impairment, communication activity limitation, and communication participation restriction. The CAPAL has been shown to be a valid and reliable tool. In short, content validity was established using focus groups composed of alaryngeal speakers (n=15), their relatives (n=24), and speech therapists (n=10). They were asked to finalize the appropriate items to be included in the final version. Construct validity was established by comparing CAPAL results with 2 validated self-administered questionnaires: the Chinese version of the Medical Outcomes Study 36-item Short-Form Instrument and the Chinese-Cantonese version of the Hospital Anxiety and Depression Scale using Pearson r correlation coefficients. Significant correlations that varied from 0.26 to 0.6 were found between the CAPAL and the other 2 questionnaires (P < .05 for all comparisons). The internal consistency showed a Cronbach coefficient α of 0.98 and a test-retest reliability of 0.84.

Responses to each item were collected using an 11-point (0-10) EAI scale. Respondents were asked to put a cross on the point that best represented their response, with 0 corresponding to never and 10 corresponding to always.

DATA ANALYSIS

Calculation of Speech Intelligibility and Acceptability Scores

The speech intelligibility score was obtained by dividing the number of correctly transcribed words by the total number of words (220). For each individual speaker, a total of 6 speech intelligibility scores were obtained, 3 from the younger judge group and another 3 from the older judge group. The mean
Speech intelligibility scores given to each speaker by both the younger judges and the older judges were compared.

**Calculation of the CAPAL Scores**

The item scores were computed to obtain the following section scores:

1. **Self-perceived speech impairment score (1 item, maximum score = 10)**
2. **Daily communication section score (20 items, maximum score = 200)**
3. **Social communication section score (6 items, maximum score = 60)**
4. **Job section score (4 items, maximum score = 40)**
5. **Emotion section score (14 items, maximum score = 140)**
6. **Total CAPAL score (the sum of the 5 section scores) (maximum score = 450)**

Within each section of the daily communication, social communication, and job sections, items were organized in pairs; the first item addressed activity limitation, and the second item addressed participation restriction. Therefore, 2 additional scores were computed for each of these 3 sections:

1. **Communication activity limitation (CAL) scores:** the sum of scores from the first items of the paired questions within each section. The sum of CAL scores from the 3 sections resulted in the total CAL.
2. **Communication participation restriction (CPR) scores:** sum of scores from the second items of the paired questions within each section. The sum of CPR scores from the 3 sections resulted in the total CPR.

---

**RESULTS**

**PERCEPTUAL RATINGS OF SPEECH IMPAIRMENTS**

**Table 2** lists the perceptual ratings of alaryngeal speakers’ speech impairment by the 2 groups of judges. The mean ratings from younger and older judges were 7.1 and 5.3 (maximum score = 10), respectively, which are significantly different (df = 96; t = 5.25; P < .001), indicating that older judges perceived the alaryngeal speech to be less severe than younger judges.
There were also significant differences in the perceptual ratings between the different alaryngeal speaker groups within the younger and the older judge groups (1-way ANOVA: $F_{3,45} = 37.52$ [younger judges]; $F_{3,45} = 39.14$ [older judges]; $P < .001$ for both comparisons). Post hoc Scheffe tests revealed that the PD speaker group received significantly lower severity impairment ratings than the other groups of speakers ($P < .001$ for all comparisons) from both the younger and the older judges.

**SPEECH INTELLIGIBILITY SCORES OF THE CANTONESE SENTENCE INTELLIGIBILITY TEST**

Table 3 lists the speech intelligibility scores of the CSIT obtained by different speech groups. For both the younger and the older judges, there were significant differences in the CSIT scores among the alaryngeal speaker groups (1-way ANOVA: $F_{3,45} = 38.54$ [younger judges]; $F_{3,45} = 36.00$ [older judges]; $P < .001$ for both comparisons). Post hoc comparison using Scheffe tests revealed that the PD speaker group received the highest CSIT score compared with the other groups of alaryngeal speakers ($P < .001$ for all comparisons) from both group of judges.

In the orthographic transcription tasks of the CSIT, reliability was assessed by Pearson $r$ correlation coefficients. The intrajudge reliability of the older judge group was 0.74 ($P = .001$), and the interjudge reliability ranged from 0.64 to 0.72 ($P = .001$). For the younger judge group, the intrajudge reliability was 0.86 ($P = .001$), and the interjudge reliability ranged from 0.78 to 0.84 ($P = .001$).

When comparing the speech intelligibility scores obtained from the younger judges and the older judges, the younger judges demonstrated a significantly higher score than the older judges ($df = 96; t = 5.25; P < .001$).

**PERCEIVED COMMUNICATION-RELATED QOL AFTER LARYNGECTOMY**

Table 4 lists the 5 section scores and total CAPAL scores in the different alaryngeal speech groups. It should be noted that in the analysis, data in the job section were excluded because only 5% of the total participants (3 of 63) responded to the job section. Most of the participants were either retired or unemployed at the time of testing. Results of 1-way ANOVA revealed significant differences in the daily communication section score ($P = .006$), emotion section score ($P = .04$), and the total CAPAL score ($P = .01$). Electrolaryngeal speech scored the highest, whereas TE speech scored the lowest in the 4 speaker groups for these 3 scores. Post hoc Scheffe tests of these 3 scores further indicated that differences between EL and TE speech were statistically significant (daily communication section score, $P = .02$; emotion section score, $P = .04$; total CAPAL score, $P = .02$).

Table 5 lists the CAL scores of different speech groups. Results of 1-way ANOVA revealed significant differences in daily communication ($P < .01$) and total score ($P < .01$) among different alaryngeal speaker groups. Electrolaryngeal speech scored the highest, whereas TE speech scored the lowest among the 4 speaker groups. Post hoc Scheffe tests of these 3 scores further indicated that differences between EL and TE speech were statistically significant ($P < .01$).

Table 6 lists the CPR scores. Significant differences were found with daily communication ($P < .01$) and total scores ($P < .01$) among the 4 alaryngeal speaker groups. Electrolaryngeal speech scored the highest, whereas PD speech scored the lowest among the 4 speaker groups. Post hoc Scheffe tests of these 3 scores further showed that differences between EL and PD speech were statistically significant ($P < .01$).
One of our objectives was to compare the speech intelligibility and acceptability among different alaryngeal speaker groups. Results revealed that PD speakers had the highest speech intelligibility and acceptability scores compared with those using the other 3 alaryngeal speech methods. This finding is consistent with that reported from studies on lexical tone production among Cantonese-speaking alaryngeal speakers. Ching et al found that PD speakers had the highest lexical tone intelligibility, whereas Yiu et al also showed that both PD and TE speakers were more efficient in conveying lexical tone. The present findings provide further support to the contention that lexical tone production is critical in determining the speech intelligibility of Cantonese-speaking alaryngeal speakers.

In comparing the speech intelligibility scores of CSIT obtained from Cantonese alaryngeal speakers in the present study with the scores obtained from other nontonal language studies using the AIDS, the Cantonese-speaking alaryngeal speakers as a whole seemed to show relatively lower speech intelligibility.

**Table 4. Communication Activity and Participation After Laryngectomy (CAPAL) Questionnaire Scores of Alaryngeal Speakers**

<table>
<thead>
<tr>
<th>Questionnaire Section</th>
<th>Score According to Alaryngeal Speech Mode</th>
<th>F&lt;sub&gt;4,45&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-perceived speech impairment, max score = 10</td>
<td>ES 2.00 (2.24) EL 2.93 (2.70) TE 2.08 (1.32) PD 2.93 (2.52)</td>
<td>0.59</td>
</tr>
<tr>
<td>Daily communication, max score = 200</td>
<td>ES 59.00 (32.92) EL 81.86 (45.63) TE 37.31 (33.31) PD 42.00 (22.23)</td>
<td>4.68</td>
</tr>
<tr>
<td>Social communication, max score = 60</td>
<td>ES 23.14 (13.45) EL 24.14 (17.97) TE 10.46 (16.52) PD 18.67 (18.62)</td>
<td>1.61</td>
</tr>
<tr>
<td>Job, max score = 40</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Emotion, max score = 140</td>
<td>ES 23.00 (24.69) EL 38.36 (28.17) TE 11.85 (17.81) PD 29.20 (18.87)</td>
<td>3.27</td>
</tr>
<tr>
<td>Total, max score = 450</td>
<td>ES 107.14 (67.78) EL 164.29 (83.87) TE 72.70 (57.71) PD 92.80 (54.22)</td>
<td>3.98</td>
</tr>
</tbody>
</table>

Abbreviations: EL, electrolarynx; ES, esophageal; max, maximum; NA, not available; PD, pneumatic device; TE, tracheosophageal.

Data are given as means (SDs) except where indicated.

P<.01.

There is no standard deviation for this score.

Analysis of variance result for the job score was not available owing to the limited number of participants.

**Table 5. Communication Activity Limitation (CAL) Scores of Alaryngeal Speakers**

<table>
<thead>
<tr>
<th>Questionnaire Section</th>
<th>Score According to Alaryngeal Speech Mode</th>
<th>F&lt;sub&gt;4,45&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily communication, max score = 100</td>
<td>ES 31.86 (17.78) EL 40.71 (22.99) TE 17.69 (12.11) PD 23.27 (10.91)</td>
<td>5.04</td>
</tr>
<tr>
<td>Social communication, max score = 30</td>
<td>ES 11.71 (6.85) EL 12.07 (9.08) TE 5.23 (8.26) PD 11.67 (9.78)</td>
<td>1.79</td>
</tr>
<tr>
<td>Job, max score = 20</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Total, max score = 150</td>
<td>ES 43.57 (24.54) EL 64.78 (28.39) TE 27.92 (18.66) PD 34.93 (17.83)</td>
<td>4.35</td>
</tr>
</tbody>
</table>

Abbreviations: EL, electrolarynx; ES, esophageal; max, maximum; NA, not available; PD, pneumatic device; TE, tracheosophageal.

Data are given as means (SDs) except where indicated.

P<.01.

There is no standard deviation for this score.

Analysis of variance result for the job section score was not available owing to the limited number of participants.

**Table 6. Communication Participation Restriction Scores of Alaryngeal Speakers**

<table>
<thead>
<tr>
<th>Questionnaire Section</th>
<th>Score According to Alaryngeal Speech Mode</th>
<th>F&lt;sub&gt;4,45&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily communication, max score = 100</td>
<td>ES 27.14 (16.00) EL 41.14 (22.42) TE 20.62 (24.23) PD 18.73 (13.99)</td>
<td>3.48</td>
</tr>
<tr>
<td>Social communication, max score = 30</td>
<td>ES 11.43 (6.75) EL 12.07 (9.03) TE 5.23 (8.34) PD 7.00 (9.58)</td>
<td>1.78</td>
</tr>
<tr>
<td>Job, max score = 20</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Total, max score = 150</td>
<td>ES 38.57 (22.59) EL 58.21 (32.02) TE 30.85 (33.99) PD 25.73 (20.16)</td>
<td>3.00</td>
</tr>
</tbody>
</table>

Abbreviations: EL, electrolarynx; ES, esophageal; max, maximum; NA, not available; PD, pneumatic device; TE, tracheosophageal.

Data are given as means (SDs) except where indicated.

P<.05.

There is no standard deviation for this score.

Analysis of variance result for job section score was not available owing to the limited number of participants.

**SPEECH INTELLIGIBILITY AND ACCEPTABILITY OF ALARYNGEAL SPEECH MODES**

One of our objectives was to compare the speech intelligibility and acceptability among different alaryngeal speaker groups. Results revealed that PD speakers had the highest speech intelligibility and acceptability scores compared with those using the other 3 alaryngeal speech methods. This finding is consistent with that reported from studies on lexical tone production among Cantonese-speaking alaryngeal speakers. Ching et al found that PD speakers had the highest lexical tone intelligibility, whereas Yiu et al also showed that both PD and TE speakers were more efficient in conveying lexical tone. The present findings provide further support to the contention that lexical tone production is critical in determining the speech intelligibility of Cantonese-speaking alaryngeal speakers.

In comparing the speech intelligibility scores of CSIT obtained from Cantonese alaryngeal speakers in the present study with the scores obtained from other nontonal language studies using the AIDS, the Cantonese-speaking alaryngeal speakers as a whole seemed to show relatively lower speech intelligibility.
Bridges\textsuperscript{6} reported a mean sentence intelligibility score of 87.69\% for English-speaking TE speakers and 81.71\% for the ES speakers with naive listeners. A higher intelligibility score (97.4\%) was reported by McAuliffe et al\textsuperscript{22} with a group of TE speakers. One possible explanation for the discrepancy between the Cantonese- and English-speaking alaryngeal speakers might have been the nature of the lexical tone in Cantonese. Cantonese alaryngeal speakers would have to control the lexical tone variations in order to convey the verbal messages effectively. With an inability or difficulty in producing the lexical tone in some alaryngeal speech, the intelligibility would thus be lower. In fact, it has been reported that different alaryngeal speech methods are different in conveying lexical tones.\textsuperscript{11,14}

Interestingly, younger listeners tended to understand alaryngeal speech relatively better with a lower acceptability toward it. The present findings are similar to those reported by Clark,\textsuperscript{10} who found that age of the listeners has a potential influence on the acceptability and understanding of alaryngeal speech.

**COMMUNICATION-RELATED QOL OF ALARYNGEAL SPEECH MODES**

Another objective of this study was to compare the communication-related QOL among speakers of the 4 alaryngeal speech modes. Although the PD speakers received the best speech intelligibility ratings, they did not perceive themselves as having the best communication-related QOL, as revealed by the CAPAL scores. The present results reveal that alaryngeal speakers who used TE speech methods received the lowest CAPAL scores (ie, the best communication-related QOL), followed by the PD speech. However, such differences between TE and PD speech were not statistically significant (\(P > .05\)). The use of an external visible device in the PD speech method may have made it cosmetically less favorable than TE speech mode. Such a factor may have further imposed on the patient's self-perception of communication efficiency, hence resulting in relatively poorer QOL. The present findings support the findings of the QOL studies, which used the ICF\textsuperscript{19} framework. According to the ICF,\textsuperscript{19} an impairment does not necessarily result in similar fashion of activity limitation and participation restriction. Therefore, causal relationship among the 3 levels should not be assumed. In fact, it has been advocated that speech impairment in alaryngeal speakers does not necessarily reflect the extent of communication activity limitation and participation restriction.\textsuperscript{23} This is borne out by our findings.

**CONCLUSION**

Findings from the present study shed light on the management of the communication-related QOL in alaryngeal speakers in general, although the data were obtained from the Chinese population. On the one hand, this study found that listeners of different ages had a considerable difference in their abilities in understanding and accepting alaryngeal speech. The younger the listeners, the easier it was for them to understand alaryngeal speech, but they had a lower level of acceptability toward it. On the other hand, the older the listeners, the more difficult it was for them to understand alaryngeal speech, but they had a better level of acceptability toward it. This finding highlights the need to consider the ages of the alaryngeal speaker's significant others, relatives, and friends. In the clinical study of laryngectomy, the age of the listeners or judges in rating alaryngeal speech should be taken into consideration.

Moreover, in postlaryngectomy speech rehabilitation, the choice of different alaryngeal speech methods for new alaryngeal speakers should be based on empirical data. The present results indicate that ES speakers present with the poorest speech proficiency, and EL speakers perceive the poorest communication-related QOL. Although PD speech, which is more popular in the Chinese-speaking community, seems to be superior owing to its higher speech intelligibility in Chinese and listeners' acceptability ratings, TE speech reveals relatively better overall communication-related QOL than PD speech. Therefore, both TE and PD speech have their own advantages in the Chinese population. Considering that the ultimate goal in postlaryngectomy speech rehabilitation is to enhance an alaryngeal speaker's QOL,\textsuperscript{23} we contend that TE speech achieves an overall more favorable outcome and QOL for Chinese alaryngeal speakers. We also argue that this conclusion—TE speech achieves a better outcome—can be applied equally well to Western communities.

Submitted for Publication: January 27, 2008; final revision received May 12, 2008; accepted May 23, 2008.

Correspondence: Edwin M.-L. Yiu, PhD, Voice Research Laboratory, Division of Speech and Hearing Sciences, University of Hong Kong, 5/F Prince Philip Dental Hospital, 34 Hospital Rd, Hong Kong (eyiu@hku.hk).

Author Contributions: All of the authors had full access to all the data in the study and take full responsibility for the integrity of the data and the accuracy of the data analysis. **Study concept and design:** Law and Yiu. **Acquisition of data:** Law. **Analysis and interpretation of data:** Law and Ma. **Drafting of the manuscript:** Law and Ma. **Critical revision of the manuscript for important intellectual content:** Yiu. **Statistical analysis:** Ma. **Obtained funding:** Yiu. **Administrative, technical, and material support:** Law. **Study supervision:** Law and Yiu.

Financial Disclosure: None reported.

Funding/Support: This study was supported in part by a grant from the Voice Research Laboratory, University of Hong Kong (Ms Law).

Additional Contributions: The New Voice Club of Hong Kong assisted with participant recruitment.

**REFERENCES**
