Chula Med J Vol. 45 No. 11 November 2001

Intelligibility of esophageal speech in Thai laryngectomies

Nantana Pracharitpukdee*

Sirikanya Lertsaranyapong** Siripornchai Suphanakorn***

Pracharitpukdee N, Lertsaranyapong S, Suphanakorn S. Intelligibility of esophageal speech in Thai laryngectomies. Chula Med J 2001 Nov;45 (11): 951 - 62

Background : *After laryngectomy, although esophageal speech has been the most frequently advocated method of vocal rehabilitation, a number of the laryngectomized patients failed to achieve (achieve what;) because of the speech therapy time and their own attention. So a specific system for esophageal speech training to increase the intelligibility and decrease the therapy time for the Thai laryngectomees is needed.*

Objective : *To evaluate the speech intelligibility of 9 Thai laryngectomized speakers, who received speech therapy and used esophageal speech in daily living.*

Setting : *Speech and Language Pathology Unit, Department of Rehabilitation Medicine, King Chulalongkorn Memorial Hospital, Thai Red Cross Society, Bangkok 10330, Thailand*

Research design : *Confusion Matrices*

Subjects : *- 9 laryngectomized patients, 8 males and 1 female ; age range 28-74 years (mean age 52 years). These patients had received speech therapy by speech and language pathologists at the King Chulalongkorn Memorial Hospital and the Rachavithee Hospital. The speech training was systematically based on the "Esophageal voice and speech systematic training for Thai laryngectomees".*

Patients

* Speech & Language Pathology Unit, Dept. of Rehab. Med., King Chulalongkorn Memorial Hospital.

** Speech Clinic, Institute of Otolaryngology, Rajvithi Hospital

*** Department of Otolaryngology, Faculty of Medicine, Chulalongkorn University

Listeners - *a group of 6 normal hearing listeners, 1 male and 5 females, age range 20 – 38 years (mean age 28 years)*

Materials : *A list of 36 two Thai syllable words was selected from the bisyllabic lists used in audiometric test. According to this bisyllabic word list, it dose not includ of all the Thai phonemes. Some words are added using a linguistic approach, based on the stressed and unstressed syllables. Each word was printed in A4 paper, these word cards were randomly presented to the subjects.*

Methods : *(intervention / measurement / statistics)*
Audio and visual-audio recording of 36 two syllables Thai words, produced by each laryngectomees, were presented to a group of 6 normal hearing listener who orthographically transcribed their response. Listeners' responses were registered in confusion matrices and considered phoneme types.

Results : *Results indicated that the difficulty in producing the aspirated - voiceless distinction appeared in Thai esophageal speakers. For examples in the voiceless stop group included the confusions of /p/ with /ph/ (2.9%), /t/ with /th/ (3.7%), and /k/ with /kh/ (4.3%) ; and also in the voiceless affricated phoneme /c/ with /ch/ (6.8%). The voiced alveolar stop phoneme /d/ showed the lowest percentage of correct identification (73.5%). When considering the manner of production, the type of liquid had the highest number of confusions with the other phoneme class (79.1%).*

Conclusions : *The finding may be used as guideline for developing the esophageal speech training system for the Thai laryngectomees. The specific phoneme selecting system for the Thai language could increase intelligibility and reduce training time.*

Key words : *Intelligiblility, Esophageal speech, Thai laryngectomees.*

Reprint request : Pracharitpukdee N, Speech & Language Pathology Unit, Dept. of Rehab. Med., King Chulalongkorn Memorial Hospital, Bangkok 10330, Thailand.

Received for publication. May 8, 2001.

นันทนา ประชาฤทธิ์ภักดี, สิริกัญญา เลิศศรัณยพงศ์, ศิริพรชัย ศุภนคร. ความชัดเจนของเสียงพูดด้วยลมจากหลอดอาหารในผู้ป่วยไร้กล่องเสียงไทย. จุฬาลงกรณ์เวชสาร 2544 พ.ย; 45(11): 951 - 62

ปัญหา : *ถึงแม้ว่าการพูดด้วยลมจากหลอดอาหารเป็นวิธีการหนึ่งที่ผู้ไร้กล่องเสียงไทยเลือกใช้เป็นเสียงทดแทน หลังจากได้รับการผ่าตัดกล่องเสียงออกแล้ว แต่เนื่องจากการฝึกต้องใช้ระยะเวลาและความตั้งใจในการฝึกอย่างมาก ทำให้มีผู้ไร้กล่องเสียงจำนวนหนึ่งไม่สามารถฝึกการพูดจนสำเร็จได้ ดังนั้นการศึกษาความชัดเจนของการพูดด้วยลมจากหลอดอาหารจึงเป็นแนวทางสำหรับการจัดการฝึกพูดอย่างเป็นระบบเฉพาะสำหรับภาษาไทย เพื่อเพิ่มความชัดเจนของการพูดและช่วยลดระยะเวลาของการฝึกพูดด้วยวิธีนี้*

วัตถุประสงค์ : *เพื่อศึกษาความชัดเจนของการเปล่งเสียงพูดด้วยลมจากหลอดอาหารของผู้ไร้กล่องเสียงไทยจำนวน 9 ราย ซึ่งได้รับการฝึกหัดการพูดด้วยลมจากหลอดอาหาร โดยใช้ "คู่มือการฝึกระบบเสียงและการพูดสำหรับผู้ไร้กล่องเสียงไทย" จนสามารถพูดติดต่อสื่อสารได้ในชีวิตประจำวัน*

สถานที่ทำการศึกษา : *หน่วยอรรถบำบัด ฝ่ายเวชศาสตร์ฟื้นฟู โรงพยาบาลจุฬาลงกรณ์สภากาชาดไทย*

รูปแบบการวิจัย : *การศึกษาแบบสถิติเชิงพรรณา*

วิธีการ : *การเก็บรวบรวมข้อมูลแบ่งเป็น 2 ขั้นตอน คือการรวบรวมเสียงพูดของผู้ไร้กล่องเสียง โดยให้อ่านคำสองพยางค์จำนวน 36 คำพร้อมทั้งบันทึกเสียง และบันทึกวีดีทัศน์ในเวลาเดียวกัน และขั้นตอนที่ 2 คือนำแถบบันทึกเสียง และแถบบันทึกภาพ ให้ผู้ฟังที่มีการได้ยินปกติจำนวน 6 รายฟัง และเขียนถ่ายถอดคำที่ได้ยินทั้ง 36 คำลงในแบบฟอร์ม และนำคำที่รับฟังได้มาวิเคราะห์ โดยวิธี Confusion Matrices*

ประชากรที่ทำการศึกษา : *กลุ่มผู้ไร้กล่องเสียง จำนวน 9 ราย, ชาย 8 ราย หญิง 1 ราย ช่วงอายุระหว่าง 28 – 74 ปี (อายุเฉลี่ย 52 ปี) ประชากรกลุ่มนี้ได้รับการฝึกพูดด้วยลมจากหลอดอาหารจากนักอรรถบำบัด ที่โรงพยาบาลจุฬาลงกรณ์และโรงพยาบาลราชวิถี และกลุ่มผู้ฟัง เป็นผู้มีการได้ยินปกติจำนวน 6 ราย, ชาย 1 ราย หญิง 5 ราย, ช่วงอายุระหว่าง 20 – 38 ปี (อายุเฉลี่ย 28 ปี)*

ผลการศึกษา : *ผู้ไร้กล่องเสียงไทยกลุ่มนี้มีปัญหาในการเปล่งเสียง ให้มีความแตกต่างระหว่างหน่วยเสียงที่มีลมประกอบ (aspirated phonemes) กับหน่วยเสียงอโฆษะ (voiceless phonemes) ตัวอย่างเช่นเสียงกักแบบอโฆษะ (voiceless stop) /p/ ถูกแทนด้วย /ph/ (2.9%), /t/ ถูกแทนด้วย /th/ (3.7%) และ /k/ ถูกแทนด้วย /kh/ (4.3%) และหน่วยเสียงกึ่งเสียดแทรกแบบอโฆษะ (voiceless affricated phonemes) /c/ ถูกแทนด้วย /ch/ (6.8%) และหน่วยเสียงที่มีค่าความสับสนสูงที่สุด คือเสียงกักประเภทโฆษะที่มีฐานกรณ์ที่ปุ่มเหงือก (voiced alveolar stop) /d/ (73.5%) นอกจากนั้นเมื่อพิจารณาแบ่งหน่วยเสียงทั้งหมดเป็นแบบกลุ่มประเภทของการเปล่งเสียง (manner of production) พบว่าหน่วยเสียงที่มีค่าความสับสนสูงที่สุด ได้แก่ประเภทหน่วยเสียงเหลว (liquid) (79.1%)*

วิจารณ์และสรุป : *ผลการศึกษา สามารถนำไปใช้เป็นแนวทางในการพัฒนาระบบการฝึกพูดด้วยลมจากหลอดอาหารสำหรับผู้ไร้กล่องเสียงไทย เพื่อให้การฝึกเสียง และการพูดถูกต้องตามระบบเฉพาะของภาษาไทย ทำให้การพูดด้วยลมจากหลอดอาหารของผู้ไร้กล่องเสียงไทยมีความชัดเจน และลดระยะเวลาในการฝึกพูด*

After laryngectomy, one of the most important objectives is voice restoration. But the a low number of Thai laryngectomized patients (49 %) received advice from the otolaryngologists and speech and language pathologists and 13 % of these patients did not know about the voice restoration.[1] Three methods have been used to achieve voice rehabilitation in the laryngectomy: artificial larynx, esophageal voice, and tracheaesophageal voice. The artificial laryngeal consists of an external mechanism that acts like a voice source. Esophageal and tracheaesophageal speechs are similar with respect to the source vibration mechanism but the methods are dissimilar with respect to air supply.[2]

This study aimed to evaluate the speech intelligibility of Thai laryngectomized speakers, who use esophageal speech. Although this method is widely employed, it takes time for speech training : 3-12 month for English[3] and 1 - 3 months for Thai.[1] The more attention and the longer training time than other methods are barriers to acquiring the functional communication of this kind of speech. Besides, characteristically the voices of the patients who are able to use esophageal speech are limited with respect to intensity, pitch, and rate.[4] Many studies, therefore have focused on acoustic analysis and the voice characteristics. In Thailand there have been few studies on the Thai alaryngeal speech, for example, the rhythm in Thai esophageal and tracheaesophageal speech,[5] and the study on vowel length in Thai alaryngeal speech.[6] The functional communication or percentage of words correctly identified as a measure of intelligibility have not been studied with Thai speakers. The Thai consonants in initial positions, produced by esophageal patients were considered. Analysis in terms of phonetic dimensions can be useful in treatment programs because the errors would indicate specific phoneme production problems on which speech-language pathologists can focus their attention.

Materials and methods

Stimuli and subjects

A List of 36 two Thai syllable words was selected from the bisyllabic lists used in audiometric test. According to the bisyllabic word list it did not include all of the Thai phonemes. Some words are added using a linguistic approach, based on the stressed and unstressed syllables.[7] Each word was printed on A4 paper and these word cards were randomly presented to the subjects.

The list of 36 two syllables was pronounced by 9 laryngectomized patients, 8 males and 1 female, age range 28 – 74 years (mean age 52 years). These patients had received speech therapy by speech and language pathologists at the King Chulalongkorn Memorial Hospital and the Rajvithi Hospital. The speech training was based on the "Esophageal voice and speech systematic training for Thai laryngectomees".[8] The data collection began when these patients were valged to be able to use esophageal speech in daily living.

The list of words produced by the laryngectomees was recorded and presented to a group of 6 normal hearing listeners, 1 male and 5 females, age range 20 – 38 years (mean age 28 years) in a word-recognition task using an open response paradigm. The errors were registered in confusion matrices and were analyzed in terms of phonetic dimensions. We assumed that the errors committed by the listeners

reflect production difficulties of the speakers. A control group of 9 normal voiced subjects produced and recorded that same list of words. The subjects were matched by sex and age to the patients.

Procedure

Recording. A National Model No. RQ. 2102 tape recorder and a Sony Model No. VCCD-TRV 11E. video recorder were used in this study. The laryngectomized patients made the recording individually in a quiet room. Both recordings were synchronized. The patients sit at the distance of 30 cm. from the tape recorder and 1.50 m. from the video recorder. The patients were instructed to read the list of words at their normal intensity with an interval of approximately 3 seconds between two consecutive words. Each word was randomly presented by the examiner behind the video recorder.

Perception task. A group of 6 normal hearing listeners carried out the perception task in the quiet room. The stimuli were presented on a Stereo Denon Model. D-08 tape recorder and on a Sony Trinitron Model. KV-G21P1 television. The perception tasks were divided into 6 times 3 stimuli each presented (randomly from the tape recorder or video recorder). Because there were 9 laryngectomees, so we had 2 series of word lists: the 9 from tape recorders and the other 9 from viedo recorders. The listeners adjusted the playback level to a comfortable intensity level, and all the other conditions were consistent for all the listeners. Before the recognition task the listeners were presented with a short self-introduction for each laryngectomee so that they could become familiar with the task. After this, the stimuli were presented, and the listeners transcribed orthographically their responses. An interval of a week was used between listening sessions in order to minimize memory effects.

Confusion Matrices

We had 2 data sets of 36 word lists for the listeners: one set from the tape recorder and the other set from the video recorder. The listeners' responses provided the data to be analyzed. We considered a response to be an error whenever the listener did not write the word that was the intended production. For example, if the intended word was / jot sak / (rank) and the listener wrote / ton sak / (teak tree), a confusion of / j / with / t / was recorded. Omissions were also scored as errors.

These errors were arranged in two separate phoneme confusion matrices. Table 1 is the confusion matrix of the esophageal speech from the audio recording, and Table 2 is the confusion matrix from the video recording. The left column represents the phonemes given as stimuli, and the rows represent the responses of the listeners. The diagonal gives the percentage of correct identifications. All the values were expressed in percentages because the number of occurrences of each phoneme was not balanced. In the right column, omission data is given. The confusion matrix from the control group of speakers is not given because of the absence of such confusions. Normal speakers had 100 % intelligibility.

Results

The most salient result shown in the confusion matrices of table 1 is that the voiceless sounds were substituted by the aspirated sounds. For instances, the bilabial voiceless stop / p / with the bilabial aspirated / ph / (2.9 %) , the alveolar voiceless stop

Table 1. Phoneme confusion matrix (in percent) in the esophageal speech from audio-recording.

R:	p	t	k	b	d	ph	th	kh	m	n	n	f	s	r	l	c	ch	j	w	?	h	omis
E:																						
p	92.6	0.6				2.9			0.3	0.2		0.7		0.7								4.32
t		85		0.9	6.2	0.3	3.7	0.6	0.3	0.4		0.4	0.9	1.1	0.9		1.2	7.7		1.9		8.6
k			83.9		1.2			4.3			1.9			0.4			1.2	3.3				8
b				90.7				1.9										1.1				3.7
d			1.2		74	0.3		0.6	0.3					0.4	3.7				0.9	1.9		11.1
ph	1.9	0.6	0.6		0.6	80.4		1.9	1.6	0.6			0.9	0.4	0.9	0.9					3.7	7.1
th			0.6				90.7														1.9	5.6
kh	0.6		3.1	0.9		1.3		76.6						0.7	0.9	0.9				1.9		9.2
m	0.6	1.2	1.2	2.8	1.2	4.5		1.2	88.1	1.6	1.9	0.4		2.2			1.2					6.9
n				0.9		0.3			1.3	85.8		0.4		1.9			0.6					8.4
n											92.6											3.7
f			0.6			0.3		2.5		0.6		85.9	0.9	1.1		0.9	0.6	0.7				9.3
s		0.6			0.6	0.5		0.6		0.2		2.6	90				0.6	0.4				6.9
r		1.9			1.9	0.3			0.5	1.6				89.8	0.9							9.3
l		1.2	0.6							0.4				0.7	74.1							13.9
c					3.7				0.3							91.7	6.8	2.2				5.6
ch												0.4					82.1	0.4	0.9			5.5
j									0.3					1.1	4.6			77.1				6.3
w																			92.6			5.5
?						1.9							0.5							94.4		
h									0.3									0.7			92.6	1.9

/ t / with the alveolar aspirated / th / (3.7%) and the velar voiceless stop / k / with the velar aspirated / kh / (4.3 %) and also the voiceless affricated sound / c / with the aspirated affricated / ch / (6.8 %)

The voiced alveolar stop / d / was the most confused phoneme (73.5 %). It was also the only voiced phoneme that was substituted with a voiceless / t / (6.2 %)

For the 3 nasal phonemes, the alveolar nasal/ n / was the most confused (85.5 %). The bilabial nasal/ m / was confused and substituted with bilabial aspirated / ph / (4.5 %) . The velar nasal / N / was not confused with other phonemes but was omitted (5.6%).

The most substituted and confusing phoneme with the other correct phonemes was the palatal semi-vowel phoneme / j / i.e., / t / with / j / (7.7 %), / k / with/ j / (3.3 %) and / c / with / j / (2.2 %)

Table 2 shows the confusion matrices of the esophageal speech from the video recorder. The characteristics of the confusion agree with the one from the audio recorder and the correct value of phonemes was higher. Only the lateral phoneme was the least corrected value (78.7 %)

Taking into account the group of the phoneme class, we considered and divided the types of phonemes into 9 groups. Table 3 shows the confusion

Table 2. Phoneme confusion matrix (in percent) in esophageal speech from audio-visual recording.

R:	p	t	k	b	d	ph	th	kh	m	n	n	f	s	r	l	c	ch	j	w	?	h	omis
E:																						
p	97.5					0.5			0.5			0.4										1.9
t		87			4.9					0.4			0.9				1.9	4.4				8.6
k			90.7	0.9	1.2			2.5						0.7				1.5		3.7		7.4
b				88.9		0.3			0.5													6.5
d		0.6			82				0.3				0.5		3.7		0.6					5.5
ph		0.6			0.6	82.6		0.6	0.8	0.2			0.5					3.8	1.9			7.9
th							94.4						0.5			0.9	0.6					3.7
kh			1.2		0.6			85.1	0.3					0.7		0.9			1.9			9.3
m	0.6	1.2		2.8	0.6	5.6		1.2	91.8	0.6		0.4		0.4				0.7				5.3
n		0.6							0.3	90.1				1.1								7.8
n											92.6											5.6
f						0.3						88.9		0.4			0.6					7.8
s												1.1	90.3					3.8				7.4
r		1.2			0.6				0.3	0.4		0.4		82.2				3.8				10.4
l			0.6		1.2	1.3		0.6		0.4				2.9	78.7							9.3
c					1.9							0.7		0.4		98.2	6.8	2.9				
ch					1.2												82.7		0.9			6.8
j				0.9											8.3			82.2				6.7
w								0.6											94.4			1.9
?							1.9				1.9									96.2		
h												0.4						3.8			98.1	1.9

Table 3. Confusion matrix (in percent) by type of phoneme in esophageal speech from audio-recording.

	Stops-VL	Stops-VD	Stops-ASP	Affricat-VL	Affricat_VD	Fricatives	Nasals	Liquids	Glides	Omissions
Stops-VL	88.1	4.8	5.1		2.5	1.1	0.3	1.9	7.9	6.3
Stops-VD	0.4	80.4	0.8				0.1	1.3	0.8	8.1
Stops-ASP	2.2	0.7	81.6	1.9		0.4	0.7	1.3		7.6
Affricat-VL		2.2		91.7	6.8		0.1		1.6	5.6
Affricat_VD					82.1	0.2			0.3	5.6
Fricatives	0.4	0.4	1.3	0.9	1.2	89.8	0.1	0.8	1.3	7.6
Nasaps	0.9	2.2	3.4		1.9	0.4	87.8	2.9		7.5
Liquides	1.1	1.1	0.2				0.2	79.1		10.6
Glides								2.1	81.5	6.1

stop-vl = voiceless stop, stop-vd = voiced stop, stop-as = aspirated stop

affric-vl = voiceless affricate, affric-vd = voiced affricate

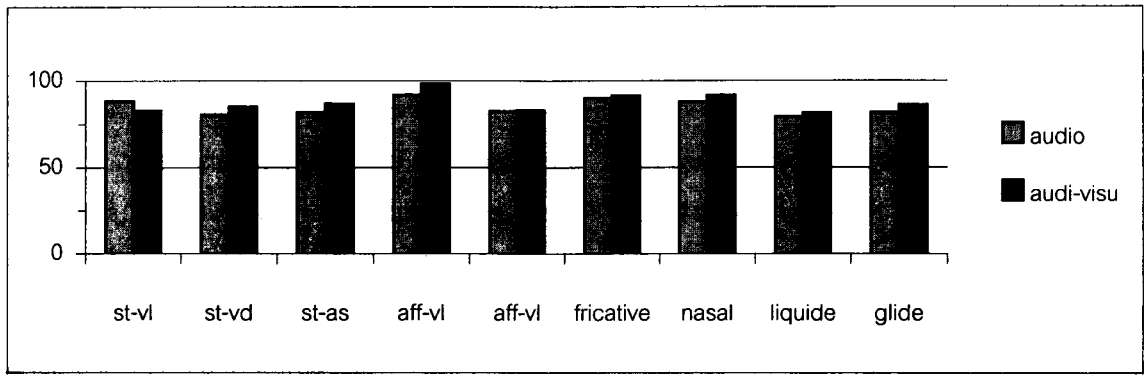
Table 4. Confusion matrix (in percent) by type of phoneme in esophageal speech from audio-visual recordering.

	Stops-VL	Stops-VD	Stops-AS	Affricat-VL	Affricat_VD	Fricatives	Nasals	Liquids	Glides	Omissions
Stops-VL	92.6	4.4	1.9		2.5	0.4	0.5	1.3	4.5	5.4
Stops-VD	0.2	84.8	0.2		0.6	0.2	0.3	1.1		5.9
Stops-AS	0.6	0.4	86.4	1.9	0.6	0.4	0.5	0.5	1.1	7.9
Affricat-VL		1.1		98.2	6.8	0.4		0.3	2.1	
Affricat_VD		0.7			82.7				0.3	6.8
Fricatives			0.2			91.1		0.3	0.3	7
Nasals	0.7	1.5	3.9			0.2	91.4	1.1	0.5	6.6
Liquides	0.6	1.1	1			0.2	0.5	81.2	0.3	10
Glides		0.4	0.2					2.4	85.7	5.3

stop-vl = voiceless stop, stop-vd = voiced stop, stop-as = aspirated stop

affric-vl = voiceless affricate, affric-vd = voiced affricate

	audio	audi - visu
st - vl	88.1	82.6
st - vd	80.4	84.8
st - as	81.6	86.4
aff - vl	91.7	98.2
aff - vl	82.1	82.7
fricative	89.8	91.1
nasal	87.8	91.4
liquide	79.1	81.2
glide	81.5	85.7



audio = audio recording, audi - visu = audio - visual recording

st - vl = voiceless stop, st - vd = voiced stop, st - as = aspirated stop

aff - vl = voiceless affricate, aff - vd = voiced affricate

Figure 1. Comparison types of phonemes from autio recording to autio-visual recording.

matrices of the manner of types of articulation from the audio recorders, and Table 4 shows the ones from the video recorders. The findings point out that the group of phonemes that were well-recognized, giving the highest scores of correct identifications was the voiceless affricated group (91.7 % ; table 3) and (98.2 % ; table 4), and that the group of liquid phoneme was the most confused (79.1 % ; table 3) and (81.2%; Table 4). The confusion values of the manner types from the audio and video recorders showed agreement: the latter was the higher score in every types.(figure 1)

Discussion

Perceptual errors allowed us to differentiate certain common characteristics of esophageal speech when considered with linguistic aspects; the specific characteristics resulting from the method of voice production. The findings show that the Thai esophageal speakers could not produce the aspirated- voiceless distinction. It is possible that at the beginning of the esophageal speech training, the patients try to bring out the amount of air from the esophagus for voice production. Therefore, If the speakers can not control the air - (articulator co-ordination), the exaggerated air spontaneously comes out. This is important because the aspirated phoneme is the specific characteristic of the Thai language, but there has not been any studies to compare. However, these two, esophageal English speakers, had a similar problem: they could not pronounce the voiced – voiceless distinction because they could not close the pharyn-goesophageal while the air stream passing. [9]

From a perceptual point of view, a portion of the errors could be considered as predictable, such as the confusions between phonemes sharing place of articulation. [10] The most difficult phoneme for the Thai esophageal speakers was the voiced alveolar stop / d /. This phoneme was confused with the voiceless alveolar stop / t / which shares the place of articulation. When there a limitation on articulation, the phonemes, which have the similar articulation placement, are selected to substitute.[11] For instance, the liquid manner type/ l / and / r,which were the most difficult type for the speakers, were substituted with other phonemes, which similar articulation placement (alveolar) such as the phoneme / l / with / t /, / n / , and / r /, and the phoneme / r / with / t / , / d / , /n /, and / l /.

It is not surprising that the study showed that the semi- vowel phoneme / w / was not confused with other phonemes and at the same time the semi-vowel phoneme / j / substituted for others phonemes with the high confusion scores. These findings agreed with the studies on English esophageal speech, which found that the vowel phonemes were well recognized. When we pronounce the vowels, the air stream passes the vocal tract without articulator closure.[2, 12]

The confusion matrices of esophageal speech showed the high score of phoneme identification (> 70 %) and the highest score was phoneme / ? / (94 % from audio recorder and 96.2 % from video recorder). It is possible that this group of speakers received speech therapy from the speech and language pathologist and that the training was also based on the Thai linguistic system. Also, the scores of the confusion matrices from the video recorder were higher than the one from the audio recorder. Thus, the intelligibility of the esophageal speech would increase when we pay attention to the esophageal speakers and look at their face.

In this study, we collected the data from the two syllable words. Sometimes the listeners could repeat the correct word but they hesitated to write it down. They thought of the meanings of the words before transcribing. This made the word different from the correct word that they perceived and repeated. This problem occurred because during the perception process, besides acoustical information, listeners make use of phonological and semantic knowledge to improve the recognition process. [13] The passage with the intended word reading is recommended for further study on intelligibility of Thai esophageal speech.

From this study we learned the characteristics of the Thai esophageal speech, and also the problems which are the cause of Thai esophageal speakers' unintelligibility. These may be a guideline for developing the effective esophageal speech system in order to increase the intelligibility and decrease the training time for the Thai laryngectomees.

Acknowledgement

This study have been supported by the Princess Mother's Medical Volunteer Foundation. The Researchers would like to thank Assissant Professor Prinya Luangpitakchumpon for his kind advice on audiometric testing.

References

1. สิริกัญญา เลิศศรัณยพงศ์, นันทนา ประชาฤทธิ์ภักดี, ศรีวิมล มโนเชี่ยวพินิจ. รายงานโครงการ / การวิจัยเรื่องบุหรี่และมะเร็งกล่องเสียง (cigarettes smoking and laryngeal cancers) สถาบันควบคุมการบริโภคยาสูบ กรมการแพทย์ กระทรวงสาธารณสุข 2543.
2. Miralles JL, Cervra T. Voice intelligibility in patients who have undergone laryngectomies. J Speech Hear Res 1995 Jun; 38(3): 564 - 71
3. Blom ED, Singer MI, Hamaker RC. A perspective study of tracheoesophageal speech. Arch Otolaryngol Head Neck Sury 1986 Apr; 112 (4): 440 – 47
4. Yingyong QI, Weinberg B. Characteristics of voicing source waveforms produced by esophageal and tracheoesophageal speakers. J Speech Hear Res 1995 Jun; 38: 536 - 48
5. Gandour J, Weinbery B, Petty SH, Dardarananda R. Rhythm in Thai esophageal speech. J Speech Hear Res 1986 Dec; 29(4): 563 - 68
6. Gandour J, Weinberg B, Petty SH, Dardarananda R. Vowel length in Thai alaryngeal speech. Folia Phonatr 1987; 39: 117 - 21
7. กาญจนา นาคสกุล. ระบบเสียงภาษาไทย. โครงการตำรา คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย. กรุงเทพฯ. โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, 2541
8. นันทนา ประชาฤทธิ์ภักดี, ศิริพรชัย ศุภนคร. คู่มือระบบการฝึกเสียงและพูดด้วยลมจากหลอดอาหารสำหรับผู้ไร้กล่องเสียงไทย. คู่มือประกอบการฝึกพูด สำหรับสมาชิกชมรมผู้ไร้กล่องเสียง โรงพยาบาลจุฬาลงกรณ์. 2540.
9. Christensen JM, Weinberg B. Vowel duration characteristics of esophageal speech. J Speech Hear Res. 1976 Dec; 19(4); 678 - 89
10. Singh S. Perceptual similarities and minimal phonetic differences. J Speech Hear Res 1971 Mar; 14(1): 113 – 24
11. พิณทิพย์ ทวยเจริญ. การวิเคราะห์ความแปรผันทางระบบเสียงในภาษาไทย. ภาษาและภาษาศาสตร์. 2540 . กค. – ธค; 16(2): 18 - 28

12. Doyle PC, Danhauer JL, Reed CG. Listeners's perceptions of consonants produced by esophageal and tracheo-esophageal talkers. J Speech Hear Disord 1988 Nov; 53(4): 400 - 7

13. Stevens KN. The quantal nature of speech, evidence from articulatory-acoustic data. In : David EE, Denes PB. (eds). Human Communication : a Unified View. New York : McGraw – hill, 1972.