



Michaelsam E. Econ, MD
Ronaldo G. Soriano, MD

Department of Otolaryngology
Head and Neck Surgery
St. Luke's Medical Center

PHONETOVOX: A Novel Prototype Device for Alaryngeal Speech

ABSTRACT

Objective: To describe a prototype improvised hand held device for alaryngeal speech.

Methods:

Design: Instrumental Innovation

Setting: Tertiary Private Hospital

Participants: Four listeners with normal hearing were native Tagalog-speakers and had no previous experience with alaryngeal speech types participated in initial trials.

Results: The prototype PHONETOVOX was fabricated using a soundproofed cellphone casing with an intra-oral sound port attachment and a cellphone was loaded with Pocket Talkbox v. 1.4.0 software. The device was tested for its ability to produce intelligible speech by using the cellphone as a substitute for the larynx using oral cavity resonators and articulators producing a synthesized sound mimicking the human voice. The PHONETOVOX produced intelligible words. Initial testing using a C-V-C Tagalog Word List had 4 listeners identify 34, 35, 47 and 54 out of 93 words (37 to 58%) with an overall average intelligibility of 46%.

Conclusion: Despite its restrictions in articulation and the wide range of results from the four listeners, our initial results may suggest the potential of PHONETOVOX as another modality for alaryngeal speech comparable to the 36 – 38 % intelligibility of commercially-available devices. Further trials with actual laryngectomees are needed to further establish intelligibility and acceptability.

Keywords: *Laryngectomy, Alaryngeal speech, Laryngeal Cancer*

Total laryngectomy is a surgical procedure involving removal of the entire larynx resulting in airway interruption and respiration being accomplished through a tracheal stoma in the lower anterior cervical area. The surgery is usually performed for advanced laryngeal carcinoma or malignancy of adjacent structures as well as severe laryngeal trauma that does not allow functional reconstruction of the organ, severe laryngeal stenosis and recurrent laryngo-tracheal papillomatosis.¹ Following total laryngectomy, the generator of vibration is essentially eliminated and patients are rendered voiceless with an altered airway. Thus, voice restoration is essential in maintaining the quality of life in these affected individuals.

Correspondence: Dr. Michaelsam E. Econ
2nd floor Department of Otolaryngology
Head and Neck Surgery
St. Luke's Medical Center
279 E. Rodriguez Ave., Quezon City 1102
Philippines
Phone: +63 2 723 0101 local 5543 / +63 0908 862 3269
Email: mikeeconmd05@gmail.com

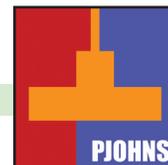
The authors declare that this represents original material that is not being considered for publication or has not been published or accepted for publication elsewhere in full or in part, in print or electronic media; that the manuscript has been read and approved by the authors, that the requirements for authorship have been met by the authors, and that the authors believe that the manuscript represents honest work.

Disclosures: The authors signed disclosures that there are no financial or other (including personal) relationships, intellectual passion, political or religious beliefs, and institutional affiliations that might lead to a conflict of interest.

Presented at the Philippine Society of Otolaryngology-Head and Neck Surgery Poster Session Contest on Surgical Innovation & Surgical Instrumentation (1st Place) December 1, 2017. Manila Hotel, Manila, Philippines.



Creative Commons (CC BY-NC-ND 4.0)
Attribution - NonCommercial - NoDerivatives 4.0 International



Alaryngeal speech modalities for voice restoration include the esophageal (SE), tracheoesophageal (TE) and electrolaryngeal (EL).²⁻⁷ Esophageal speech uses air injected into the upper esophagus that is expelled voluntarily setting the pharyngoesophageal segment into vibration thereby producing a new sound for phonation.³ Tracheoesophageal speech has the same vibratory mechanism but is achieved by creating a fistula between trachea and esophagus such as involving the use of trachea-esophageal mucosa in a one-stage procedure as described in the Amatsu technique and its modification such as the Ureta technique.^{4,5} Tracheoesophageal speech is also possible using a silicon voice prosthesis that is surgically implanted based on the principles introduced by Mozolewski in 1972 and refined by Blom-Singer in 1978.^{2,6} Electrolaryngeal speech uses an electrolarynx for speech production.⁷ Despite reported successes with tracheoesophageal speech,⁸ the surgery requires a level of mastery not all surgeons can achieve, esophageal speech may require a long learning curve on the part of laryngectomees and few master this skill,^{9,10} and the high cost of electrolaryngeal devices,¹¹ limit other alternatives for patients in our setting.

We therefore describe a new alaryngeal device as is an alternative for intelligible speech production.

METHODS

This prototype study was undertaken as proof-of-concept required for institutional review board (IRB) approval; an application for IRB approval has been made for the formal trials following this report.

A. Materials

- 360° full shell casing made from Rubber, Thermoplastic Polyurethane, Electroplated Photonic Crystals (Shenzen Super Electronic Co. Ltd, China)
- 70mm x 5mm x 2mm ethylene-vinyl acetate rubber sheet from commercially-available mouse pad (CD-R King, Philippines)
- Samsung Galaxy S7 Edge cellphone (Samsung Electronics Co. Ltd, South Korea)
- Pocket Talkbox Free APK version 1.4.0 for Android version 2.3 and up (Vito Biliti, Cayman Islands)
- Simplex[®] Mucus Extractor (Inmed Corp., Philippines)
- Cyanoacrylate adhesive glue (Mighty Bond[®], Pioneer Adhesives Inc. Philippines)
- Craft cutter (Maped, France)

B. Procedure

1. Creating a sound port and Intra-oral attachment Fabrication

A commercially-available 360° full shell casing made from rubber,

thermoplastic polyurethane and electroplated photonic crystal (Shenzen Super Electronic Co. Ltd, China) was used to provide a tight-seal over a commercially-available cellphone (Samsung Galaxy S7 Edge, Samsung Electronics Co. Ltd., Gyeonggi-do, South Korea). The sound port came from the suction attachment of a commercially-available mucus extractor (Simplex[®], Inmed Corp., Quezon City, Philippines) with inner diameter of 8mm and outer diameter of 12mm. The polymerizing vinyl chloride (PVC) suction attachment and tubing were detached from the collection chamber. (Figure 1A) The portion where the suction attachment tapers into the tubing was cut using the cutter. (Figure 1B) The cut end of the suction attachment was inserted into the preformed speaker hole on the phone casing. Cyanoacrylate adhesive glue was used to hold it in place. (Figure 1C). The remaining flexible PVC tubing which was removed from the suction attachment with length 380 mm, outer diameter 4 mm, inner diameter 3 mm and thickness of 1 mm was used as an intraoral device. The cut end was inserted into the sound port while the transparent end was meant to be placed into a user's mouth. (Figure 1D)

2. Soundproofing cellphone case

An ethylene-vinyl acetate rubber sheet with dimensions of 70mm x 5mm x 2mm (LxWxH) was cut from a mouse pad (CD-R King, Quezon City, Philippines) using a craft cutter (Maped, France). (Figure 2A) Cyanoacrylate adhesive glue (Mighty Bond, Pioneer Adhesives Inc., Quezon City, Philippines) was used to stick the rubber to the inner bottom compartment of the case. (Figure 2B)

3. Downloading and using the software

The Pocket Talkbox Free APK Version 1.4.0 for Android version 2.3 and up (Vito Biliti, Cayman Islands) was downloaded from Google Play Store (https://play.google.com/store/apps/details?id=com.Bright_Blue_LED.PocketTalkbox) and installed (September 5, 2017). The program was opened and **MAJ tab** was pressed. The **F musical key tab** (violet) was used during articulation. (Figure 3)

4. Use of the device and proper articulation technique

The intra-oral attachment was held in place towards the oropharynx in the midline over the posterior third of the tongue by each author. (Figure 4A) Lip, tongue, mouth and jaw movements were exaggerated during articulation to produce sound. (Figure 4B)

C. Pre-testing the device

Volunteers with normal hearing on pure tone audiometry and no previous experience with alaryngeal speech types were recruited from personal and professional acquaintances to participate in the initial

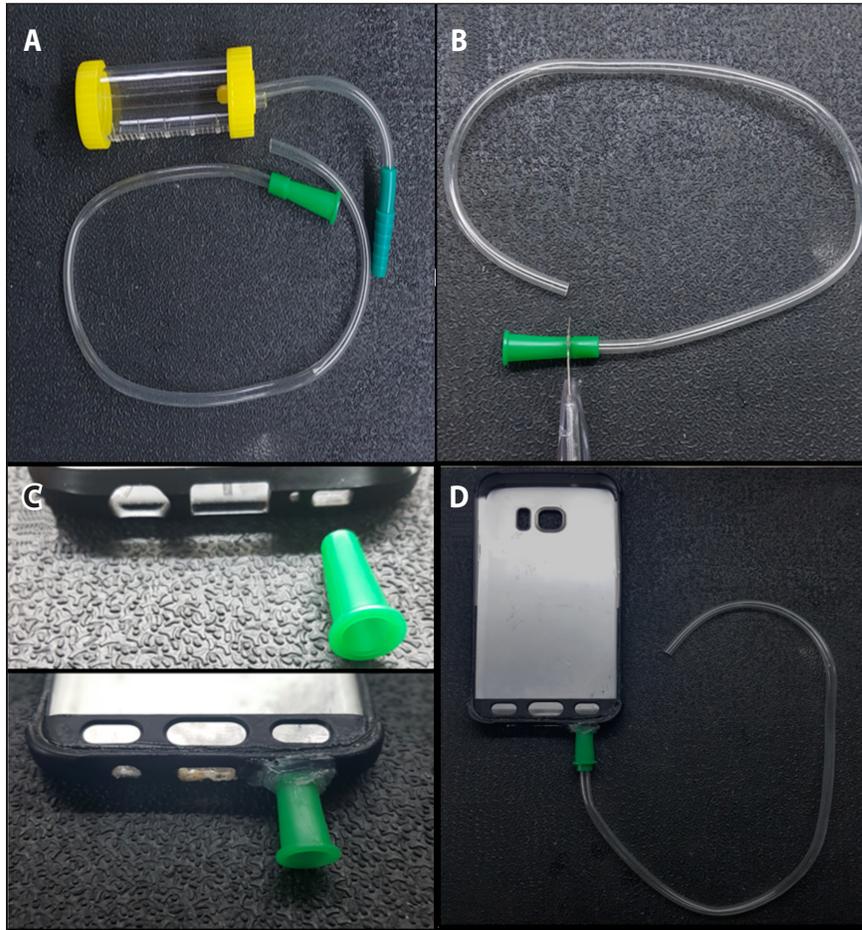


Figure 1 **A.** Simplex® Mucus Extractor, detached suction attachment and PVC tubing; **B.** Cutting suction attachment (**Sound port**) from tubing (**Intra-oral attachment**); **C.** Attaching sound port using cyanoacrylate glue; **D.** **Phonetovox** (360 full shell casing with soundport and intra-oral attachment).

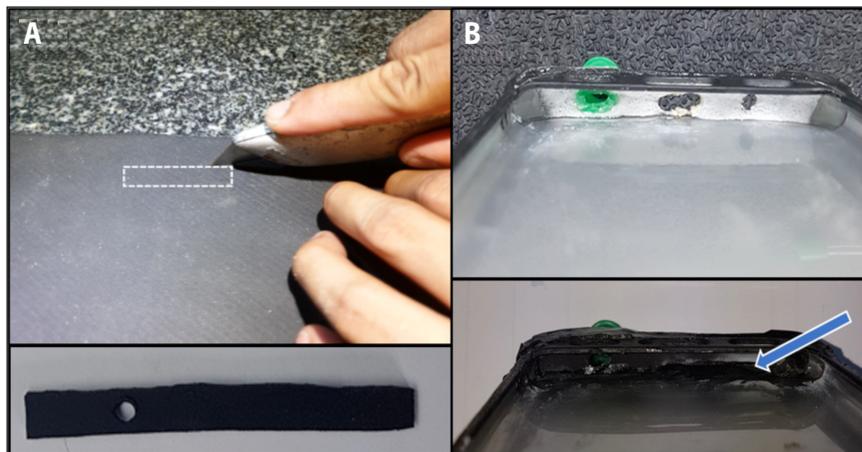


Figure 2 **A.** Using a craft cutter, a 70mm x 5mm x 2mm EVA rubber sheet was taken from a mouse pad; **B.** Sound-proofing; arrow points to EVA rubber sheet glued to inner bottom compartment of case.

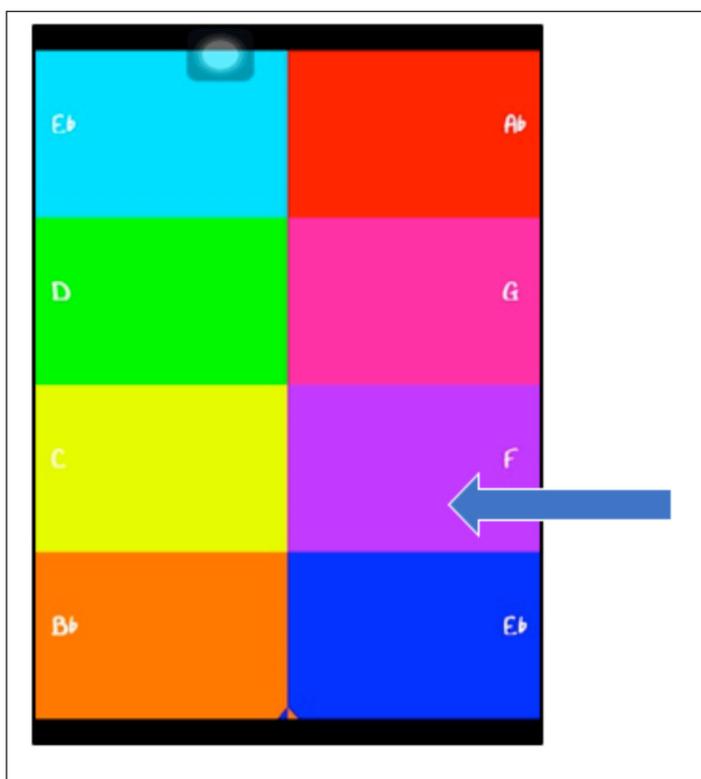


Figure 3. Keynote panel of Pocket Talkbox (Free); arrow points to F (violet) panel used

trials. All listeners had to be native Tagalog speakers and possess at least a high school education. Informed consent was obtained from volunteers meeting these criteria. The authors tested the intelligibility of the device using the speech material Consonant-Vowel-Consonant Tagalog Word List in a quiet room.¹² The listeners were blinded to the words comprising the word list. One author used the PHONETOVOX to articulate each word from the list two times. Each listener was instructed to not look at the speaker and list the intelligible words that they heard, using a pen and a blank piece of paper.

D. Data Analysis

Data analysis was accomplished using simple descriptive statistics. The testing author checked the written lists against the actual word list and tabulated the raw scores and corresponding percentages. Results were compared with values obtained from a review of related literature.

RESULTS

The prototype PHONETOVOX was successfully fabricated using a soundproofed cellphone casing with an intra-oral sound port attachment. The cellphone was easily loaded with Pocket Talkbox v. 1.4.0 software. The device utilized the sound produced by the cellphone as a substitute for the larynx, transmitting sounds from the oral cavity resonators and articulators, producing a synthesized sound mimicking the human voice.

Four listeners (3 female and 1 male) aged 26 to 29 (mean age 27.8) with normal hearing on pure tone audiometry and no previous experience with alaryngeal speech types participated in the initial trials. All listeners possessed at least a college education and were all native speakers of Tagalog.

The PHONETOVOX produced intelligible words. Initial testing using a C-V-C 93 item Tagalog word list had 4 listeners identify 34, 35, 47 and 54 out of 93 words (37 to 58%) with an overall average intelligibility of 46%. (Table 1)

DISCUSSION

The prototype PHONETOVOX /fonetovox/, which means “phone to voice” is an improvised device for alaryngeal speech fabricated from a modified cell phone casing reinforced with EVA rubber sheet for soundproofing and connected to a PVC sound port and intraoral tubing.

The concept behind phonetovox is different from other aforementioned modalities of alaryngeal speech because it utilizes the principles of a Talkbox. A Talkbox is a sound-effect gadget that allows musicians to modify the sound of their musical instrument by

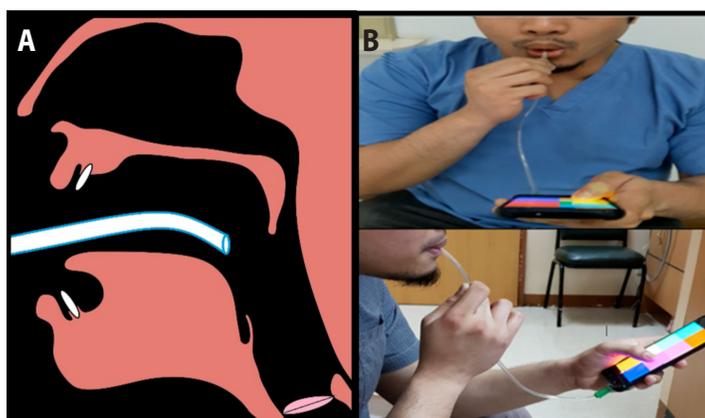


Figure 4. A. Positioning of intra-oral attachment inside oral cavity; B. Pre-testing the PHONETOVOX with proper placement of intra-oral attachment and handling of device.

Table 1. Initial PHONTOVOX trial among 4 listeners using a 93-item C-V-C word list

PARTICIPANTS	SCORE (n out of 93)	PERCENTAGE (%)
Listener A	54	58
Listener B	47	51
Listener C	35	38
Listener D	34	37
AVERAGE	42.5	46

remodeling the resonance of the sound.¹³ The phonetovox utilizes commercially-available Pocket Talkbox Ver 1.4.0 software to produce a monotone sound but other musical software such as MorphWiz-Play Ver. 1.1 software (Jordan Ruddess, U.S.A) downloaded from (<http://www.wizdommusic.com>), or your own digitally synthesized recorded sounds may be used. Substituting the eliminated sound-generating organ with a monotone sound produced by the cellphone and using the capacity of the oral cavity to alter sound to produce speech is what makes phonetovox unique from other types of alaryngeal speech.¹⁴

In constructing the device, the case should fit tightly to promote soundproofing and fullness of sound coming out of the sound port. This is achieved by lining a rubber sheet around the casing's speaker hole.¹⁵ Holding the tube in place, avoiding contact with the tongue or mucosa of the oral cavity, pooling of saliva inside the tube, and synchronization of articulation with key note tab pressing minimizes distortion of the sound leading to unnecessary noise production and unintelligibility of sound. Slow articulation with exaggeration of lip, tongue, mouth, jaw movements during articulation may enhance word intelligibility. Vowel articulation may be achieved with forward movement and high tongue positions.¹¹ The authors identified that the more polysyllabic a word the more intelligible it may become. Monosyllabic words in Phonetovox speech are unintelligible (e.g. *tren, zoo, nars*).

In a study by Weiss *et al.*, the intelligibility of two transcervical electrolarynges, Model 5 (Western Electric Company, New York, U.S.A) and Servox (Siemens, Berlin, Germany) was compared using a 66-item English word list comprising primarily of CVC syllables.¹⁶ The talker intelligibility was assessed from six laryngectomized men. The highest score for the Western Electric prosthesis was 54% and the lowest was 16%. The scores for the Servox were 55% and 19%, respectively. The overall intelligibility range scores were 38% and 36%, respectively.¹⁶ In another study by Sleeth the overall intelligibility of tracheoesophageal speech of 15 laryngectomy patients using Modified Rhyme's Test was 71% with most intelligible speaker at 89%, and the least intelligible at 54%.⁸ In our study, the overall intelligibility of Phonetovox was 46%, with highest score for intelligibility at 58% and the lowest at 37%.

Despite PHONETOVOX showing a higher intelligibility over both electrolarynges in the study of Weiss *et al.*, our trial is not comparable to the previous study since we tested the device on a non-laryngectomized individual. Thus, we cannot rule out the contribution of the speaker's intact larynx to the speech produced. Other limitations of our trial include lack of a standard comparator, testing in an unblinded manner, and only using a Tagalog word list for intelligibility testing.

The authors recommend actual trials among total laryngectomees to establish the intelligibility and acceptability of PHONETOVOX and in comparison with other types of alaryngeal speech. Our IRB application

for such a trial is in process. Other applications of the device such as use in tracheostomy patients may also be investigated later.

In conclusion, our initial results may suggest the potential of PHONETOVOX as another modality for alaryngeal speech. Despite its restrictions in articulation and the wide range of results from the four listeners, it may be comparable to the 36 – 38 % intelligibility of commercially-available devices.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. William Lim and Dr. Ray Casile for sharing their expertise in the field of Laryngology and Mrs. Sheryl Sibug -Wong for sharing her expertise in the field of Speech and Language Pathology.

REFERENCES

- Ceacir O, Hainarosie R, Zainea V. Total laryngectomy – past, present, future. *Maedica – a Journal of Clinical Medicine*. [Internet] 2014 Jun 18; 9(2): 210-216. [cited 2017 Nov 17]; Available from: https://docksci.com/total-laryngectomy-past-present-future_5a626073d64ab25013ae9f3e.html.
- Lorenz KJ. Rehabilitation after Total Laryngectomy—A Tribute to the Pioneers of Voice Restoration in the Last Two Centuries. *Frontiers in Medicine* [Internet] 2017 June; 4:81. [cited 2018 Oct 25]; Available from: <https://www.frontiersin.org/articles/10.3389/fmed.2017.00081/full>.
- Kapila M, Deore N1, Palav RS, Kazi RA, Shah RP, Jagade MV. A brief review of voice restoration following total laryngectomy. *Indian Journal of Cancer*. [Internet] 2011 Jan-Mar;48(1):99-104. [cited 2018 Oct 25]; Available from: <https://www.ncbi.nlm.nih.gov/pubmed/21248437>.
- Vieira M, Maia A, Ribeiro J. Speech rehabilitation after laryngectomy with the amatsu tracheoesophageal shunt. *Auris, nasus, larynx*. [Internet] 1999 February; 26. 69-77. [cited 2018 Oct 25]; Available from: https://www.researchgate.net/publication/13214190_Speech_rehabilitation_after_laryngectomy_with_the_amatsu_tracheoesophageal_shunt.
- Ureta C. V., Go C. E. Primary Voice Reconstruction in Total Laryngectomy Using The Funnel Technique. *Philippine Journal of Otolaryngology – Head and Neck Surgery*. [Internet] 1998; 13:52-58. [cited 2018 Oct 25]; Available from: <https://journal.pso-hns.org/2015/05/07/vol-13-no-1-1998/>.
- Blom, E. D. Tracheoesophageal Voice Restoration: Origin – Evolution – State-of-the-Art. *Folia Phoniatrica et Logopaedica*. [Internet] 1999; 52(1-3), 14–23[cited 2018 Oct 25]; Available from: <https://www.karger.com/Article/Pdf/21508>.
- Liu H, Ng M. Electrolarynx in voice rehabilitation. *Auris, nasus, larynx*. [Internet] 2007 January; 34. 327-32. [cited 2018 Oct 25]; Available from: https://www.researchgate.net/publication/6562618_Electrolarynx_in_voice_rehabilitation.
- Sleeth LE. Exploring intelligibility in tracheoesophageal speech: a descriptive analysis. [Electronic Thesis and Dissertation Repository]. The University of Western Ontario. [Internet] 2012 Jun. [cited 2017 Nov 16] Available from: <https://www.researchgate.net/publication/266853779>.
- Hillman R, Walsh M, Wolf G, Fisher S, Hong W. Functional outcomes following treatment for advanced laryngeal cancer. Part 1. Voice reservation in advanced laryngeal cancer. Part II. Laryngectomy rehabilitation: the state-of-the-art in the VA system. *Ann Otol Rhinol Laryngol*. [Internet] 1998; 107(Suppl 172, Pt 2):1–27. [cited 2018 Oct 25]; Available from: <https://www.ncbi.nlm.nih.gov/pubmed/9597955#>.
- Kresić S, Veselinović M, Mumović G, Mitrović SM. Possible factors of success in teaching esophageal speech. *Med Pregl* [Internet] 2015; LXVIII (1-2): 5-9. [cited 2018 Oct 25]; Available from: <https://www.ncbi.nlm.nih.gov/pubmed/26012237>.
- Maruthy S, Mallet MK, Bellur R. Comparison of esophageal and tracheoesophageal speech modes in dual-mode alaryngeal speakers. *J Laryngol Voice* [Internet] 2014 4(1); 6-11. [cited 2017 Oct 4]; Available from: <http://www.laryngologyandvoice.org/text.asp?2014/4/1/6/141444>.
- Lim RK, Bernhardt BM, Stemberger JP. University of British Columbia Cross-linguistic project, Tagalog words list 2015. [Internet]. [cited 2017 Sep 25] Available from: http://phonodevelopment.sites.olt.ubc.ca/tagalog-word-lists_basic_extension-a_extension-b_mar2015/.
- McCarron B. The Talkbox FAQ [Internet] c1996-2014 [cited 2017 September 3] Available from: <http://www.blamepro.com/talkbox.htm>.
- Mcgrath, J. How Talkboxes work. HowStuffworks [Internet]. [updated 2011 Jun 27; [cited 2017 Sep 3] Available from: <http://electronics.howstuffworks.com/gadgets/audio-music/talk-box.htm>.
- Woodford C. Soundproofing. Explain that stuff! [Internet] c2009-2016 [Last updated 2016 Dec 18; [cited 2017 Nov 15] Available from: <http://www.explainthatstuff.com/soundproofing.html>
- Weiss MS, Basili AG. Electrolaryngeal speech produced by laryngectomized subjects: perceptual characteristics. *J Speech Hear Res*. [Internet] 1985 Jun [cited 2017 Nov 17]. 28(2): 294-300. Available from: <https://jshlr.pubs.asha.org/article.aspx?articleid=1778128>.