

2012-04-16

The Presence of Regional Accents in Electrolarynx Speech and the Resultant Effect on Overall Intelligibility.

Brian Madden

Technological University Dublin, brian.madden@tudublin.ie

Eugene Coyle

Technological University Dublin, Eugene.Coyle@tudublin.ie

Follow this and additional works at: <https://arrow.tudublin.ie/teapotcon>



Part of the [Biomedical Commons](#), [Biomedical Devices and Instrumentation Commons](#), [Computer Engineering Commons](#), [Other Biomedical Engineering and Bioengineering Commons](#), and the [Signal Processing Commons](#)

Recommended Citation

Madden, B., Coyle, E.: The Presence of Regional Accents in Electrolarynx Speech and the Resultant Effect on Overall Intelligibility. *Signals and Systems Conference (ISSC 2012)*, 2012, NUI Maynooth, Ireland. doi: 10.1049/ic.2012.0213

This Conference Paper is brought to you for free and open access by the tPOT: People Oriented Technology at ARROW@TU Dublin. It has been accepted for inclusion in Conference Papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, gerard.connolly@tudublin.ie, vera.kilshaw@tudublin.ie.

Funder: ABBEST

The presence of regional accents in electrolarynx speech and the resultant effect on overall intelligibility.

Brian Madden^{*}, Eugene Coyle^{}**

** Department of Electrical Engineering Systems,
Dublin Institute of Technology, Dublin,
email: brian.madden@dit.ie*

***Visiting Fulbright Scholar,
Global Policy Research Institute,
Purdue University, Indiana.
email: coylee@purdue.edu*

During voiced speech, the larynx provides quasi-periodic acoustic excitation of the vocal tract. In most electrolarynxes, mechanical vibrations are produced by a linear electromechanical actuator, the armature of which percusses against a metal or plastic plate at a frequency within the range of glottal excitation.

In this paper, a phonological analysis of a section of results from an online perceptual intelligibility test was performed which compared speech produced using a novel hands-free electrolarynx and a commercially available electrolarynx. A portion of the test consisted of a closed-set format containing a selection of four sets of four random CVC audio samples (recorded by two speakers - 1 male, 1 female - using the Servox™ and the hands-free pager motor design). Each survey participant was requested to listen to every recording and then choose the word they thought most closely resembled the recording in which they heard.

The phenomenon referred to as Irish-English, as documented by Hickey [1], highlights the historical development of the English language and how its pronunciation currently varies throughout the country. The two speakers used in the intelligibility test originated from two phonologically contrasting regions – a male from the East and a female from the West. These differences are analysed with the aid of findings by Hickey and assessed as to whether they could potentially improve or hinder the intelligibility of an utterance.

Keywords – larynx, electro-larynx, intelligibility, phonology, pronunciation.

I INTRODUCTION

The larynx, located below the throat, is a small organ, which is attached to the windpipe. It performs three main functions; it acts like a valve that can close, to prevent solids or liquids from entering the airways during swallowing; it allows air, which is inhaled through the nose and mouth to reach the lungs; and it contains the vocal cords, which vibrate when air is exhaled through them. This, in essence, is how one speaks. Laryngeal cancer occurs when malignant cells form in the tissues of the larynx. When a person undergoes a total laryngectomy, a surgical procedure involving the complete removal of the larynx and the redirection of the airway through a sizeable hole in the throat called a stoma – as shown in figure 1-,

speech will never be achievable as it previously once was.

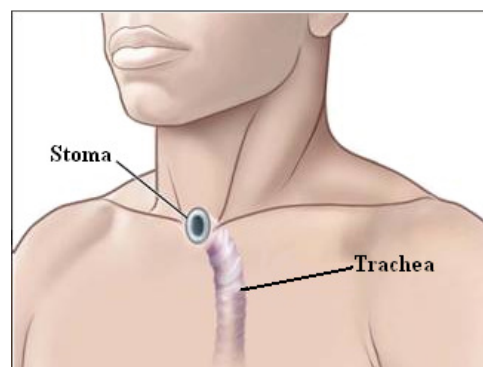


Figure 1: Redirection of airways following a TL [2].

a) Speech rehabilitation methods

The methods used to restore a laryngectomee's ability to speak are classified into three main groups: oesophageal speech, tracheo-oesophageal speech, and electrolarynx speech. Oesophageal speech is performed by driving air from the mouth into the oesophagus to inflate it and then forcing it back out to produce a burp-like sound. As in normal speech, the articulators of the upper vocal tract (the mouth, lips, tongue, nose etc.) are used to shape the sound into speech. The length of utterance is limited by the volume of air that can be held in the oesophagus compared to the lungs, during a single utterance in normal speech. In tracheo-oesophageal speech, air from the lungs is transferred from the trachea to the oesophagus via a surgically inserted one-way air valve, often referred to as a voice prosthesis. The sound produced is often similar to oesophageal speech, but the laryngectomee can speak with a more normal breathing pattern. During tracheo-oesophageal speech, a finger must usually be held over the stoma to direct air into the oesophagus, rather than letting it escape through the stoma.

Laryngectomy patients also use an electrolarynx to facilitate speech following their surgery. Harold Barney in Bell Laboratories, New York, originally invented the device in the late 1950s and its design has remained largely unchanged since. The device is pressed against the surface of the neck to generate an audible speech sound from the mouth providing a relatively intelligible means of communication for patients. However, it has inherent drawbacks such as the need for a free hand to operate the device and a difficulty experienced by many laryngectomy patients in adapting to use it. The tone and quality of the electrolarynx is often deemed unappealing to a person with a normal speaking voice. However, due to the urgent need for a laryngectomee to communicate, the device is often reluctantly accepted.

Added to these rudimentary limitations, electrolarynx speech also suffers from a number of additional constraints which include the existence of extraneous noise, due to the leaking of acoustic energy emanating from the percussive actuator; the low-frequency spectral deficiency, due to the attenuation of lower harmonics during the transmission through the neck tissue; and the artificial quality present, due to the devices' fixed frequency and amplitude.

Research to date has focused on the improvement of the quality of speech produced by the electro-larynx. Some significant contributions have been made by Houston et al. [3] who developed an electro-larynx which used digital

signal processing to create a superior quality of sound. Shoureshi et al. [4] used neural-based signal processing and smart materials to improve the sound created. Liu et al. [5] and Cole et al. [6] focused on removing the buzzing sound created by the transducer. Uemi et al [7] in 1994 developed a system that utilized measurements from air pressure that was obtained from a resistive component that was placed over the stoma to maintain the electrolarynx's fundamental frequency. Ma et al. in 1999 [8] used cepstral analysis of speech to replace the electrolarynx excitation signal with a normal speech excitation signal.

Despite all the acoustic improvements these studies have shown, they have however, been performed in isolation and have been deemed to be difficult to implement into the existing technology. Therefore, the basic concept of design which was first introduced by Barney et al. in the 1950's remains the same to this day. It has been shown that up to 50%-66% of all laryngectomees use some form of electrolarynx speech (Gray et al. [9], Hillman et al. [10]): either as a method of communication for speech rehabilitation post-surgery or as a reliable back-up in situations where esophageal or tracheo-esophageal speech is proving difficult.

b) Speech intelligibility

When determining the intelligibility of a speech signal, it is important to choose a suitable linguistic level at which to make measurements. Is it necessary to measure the accuracy at which each phonetic element is communicated in order to assess whether each word is identifiable. It is also necessary to investigate whether the communication of a sentence is clear.

This type of linguistic dismissal can introduce an additional difficulty in that individual human listeners will ultimately differ in their capability to make use of these linguistic constraints. Even though it may be deemed necessary to assess the utility of a particular channel in order to convey the meanings of real spoken utterances, listeners will inevitably vary in their capacity to comprehend the speech, depending on their own linguistic ability. Many speech intelligibility tests consist of either phonetic unit, which are composed into: nonsense syllables, words which are used in isolation or in short sentences spoken in one breath for comfort Crystal et al [11] and Mitchell et al [12].

An issue that arises through the use of nonsense syllables is that many listeners could require training in order to be able to identify the component phonetic units, and they may be confused by phonemes which don't compare well

with the spelling e.g. there, their, they're. Therefore by limiting listener reply's to real words thus allowing them to respond in ordinary spelling. This can however introduce other difficulties: firstly, that varying listeners may possess differing degrees of familiarity with the words that are being used; secondly, that some words are memorable and having heard a word once, some listeners may be biased in their usage of a particular word another time.

A possible solution to these problems includes the formulation of multiple word lists of reasonable difficulty, allowing a listener to be used within a test more than once. Another option is to create tests consisting of closed response sets, making every listener needing to make the matching choices about the word which is under test.

Egan [13] pioneered one of the first lists of words for an intelligibility test in 1948. He created the list by using the concept of "phonetic balance" which meant that the relative frequency of the phonemes in the word lists corresponded to the relative frequency of phonemes in conversational speech. He constructed 20 lists containing 50 monosyllabic words and his intention was to balance average difficulty and range of difficulty throughout the lists whilst ensuring that the phonetic units that were present were represented equally.

Monosyllabic utterances, however, can present certain difficulties when applied to electrolarynx speech as numerous studies – most notably Weiss et al. [14] and Weiss and Basili [15] – have highlighted with the confusion of voiced and unvoiced consonants by listeners in perceptual tests. Unvoiced vowels also provide an increased perceptual difficulty. .

II METHODS

a) Online intelligibility test:

To address the small number of participants that partook in the previous intelligibility tests (Madden et. 2010 [16] and Madden et al. 2011 [17]) an online survey was created which compared audio samples of the conventional electrolarynx and a novel hands-free motor pager design. The main objective of this method would be that it could potentially target a larger number of participants who could complete the survey by themselves and in their own time.

The test consisted of two parts: the first where the participants listened to two recordings of "The grandfather passage" and were asked to choose which sounded the most intelligible (one with a commercial electrolarynx and the second using a

novel pager motor design). The second test consisted of a closed-set format containing a selection of four sets of four random CVC audio samples (recorded by two speakers - 1 male, 1 female - using the Servox™ and the hands-free pager motor design). The participants were requested to listen to each recording and subsequently choose (from four possible answers) which word they thought most closely resembled the recording in which they heard.

These four groups of utterances were subsequently assessed using findings obtained from Hickey to determine whether the influence of Irish accents on electrolarynx speech altered a word so as to increase or reduce its perceived intelligibility.

b) Phonological characteristics of 'Irish-English':

Hickey reported that the spread of spoken English amongst the Irish population occurred as a result of interactions they experienced with fellow Irish (whom had regular interactions with the British). According to Hickey, mature speakers who learn a language tend to "*maintain their pronunciation of their native language and have difficulty with segments, which are unknown to them*". This resulted in the development of a version of the English language, which was heavily influenced on the pre-existing Irish language.

Examples of this would be the substitution of dental fricatives by stops (dental or occasionally alveolar – depending on the region). Another is the lenition of stops in 'Irish-English' (which differ completely from lenition of Irish words) which could be as a result of a phonological directive applied by the Irish whilst learning English. Bliss [18], reports that this occurrence is responsible for common malapropisms and the unconventional word stresses – e.g. verbs with final long vowels - which are commonly found in the 'Irish-English' language.

Therefore, in order to provide a normalised phonetic standard for the benefit of the reader, an analysis was made of these subtle differences that the Irish language had on the English language. This is outlaid in the following sections.

III RESULTS

Phonological features of utterances from test subjects:

Hickey's report separates each distinctive dialect provincially, highlighting the major differences which occur during phonation. The two speakers used in the perceptual intelligibility test originated from two distinct regions of Ireland: the male subject from Dublin (East coast) and the

female subject from Galway (West coast). Each region possesses a number of distinctly different phonological features, some of which are presented in table 4.6 and 4.7.

Fortition of dental fricatives to alveolar stops - e.g. <i>think</i> [t I ŋ k]
Front onset of /au/ - e.g. <i>town</i> [tæʊn], [t e ʊ n]
Breaking of long high vowels - e.g. <i>clean</i> [k l i j ə n]
No lowering of early modern /u/ - e.g. <i>done</i> [d ʊ n]
Glottalisation of lenited /t/ - e.g. <i>foot</i> [f ʊ t] → [f ʊ ʔ] → [f ʊ h].

Table 1: East Ireland features within 'Irish-English' [1].

Low central for /ai/ and /au/ - e.g. <i>quite</i> [k w a I t]
Shift of /tj/ to /k/ in word-internal position - e.g. <i>fortune</i> [' f ɔ r k u : n]

Table 2: West East Ireland features within 'Irish-English' [1].

These two distinct sounding dialects produced slight differences in the phonological structure of each utterance in comparison to the standard International Phonetic Association (IPA) spelling. The following tables - 3 to 6 - highlight the subtle nuances of each utterance spoken by either subject whilst using both electrolarynx devices. The most apparent difference produced by both speakers is the emphasis of the final vowels in each utterance, as noted previously by Bliss.

Word stresses in the Irish language generally occur on the first syllable of a word, except when any of the other syllables contain a long vowel - e.g. "camán" [k a m a : n] meaning "hurley". These elongations are evident with both speakers as shown in tables 3 and 4 where the vowels / o / and / a / are extended. This results in words like "Dong" [d / ɔ / ŋ] becoming "Do(on)ng" [d / ɔ / ɔ / ŋ] and "Yap" [j / æ / p] becoming [j / æ / æ / p].

Word	IPA	West	IPA
Dong	d / ɔ / ŋ	Do(ɔn)ng	d / ɔ / ɔ / ŋ
Long	l / ɔ / ŋ	Lo(ɔn)ng	l / ɔ / ɔ / ŋ
Song	s / ɔ / ŋ	"So(ɔn)ng	s / ɔ / ɔ / ŋ
Wrong	r / ɔ / ŋ	Wro(ɔn)ng	r / ɔ / ɔ / ŋ

Table 3: 'Irish-English' phonetical spelling of words with extended vowels by West of Ireland female.

Word	IPA	East	IPA
Yap	j / æ / p	Ya(a)p	j / æ / æ / p
Cap	k / æ / p	Ca(a)p	k / æ / æ / p
Trap	t / r / æ / p	Tra(a)p	t / r / æ / æ / p
Lap	l / æ / p	La(a)p	l / æ / æ / p

Table 4: 'Irish-English' phonetical spelling by East of Ireland male.

The utterances contained in table 5 and 6 illustrated characteristics of vowel breaks in closed syllables which produced an extended pause between each of the vowels - which Hickey refers to as "disyllabification of long high vowels" - as well as enhanced glottal stops. Examples of this phenomenon are "brain" [b r e i n] which becomes "brai-in" [b r e i I n], etc. in table 4 and "leap" [l / I : / p] which becomes "lee-ip" [l / I : / I / p] in table 6. This feature, more commonly found in speech originating from the East of the country (Dublin), is seen in both speakers. The female subject may have adopted this speech nuance due to having lived in Dublin for a number of years. She did, however retain characteristics from her region of origin as the low central onset of the / ai / within each of the utterances is prominent throughout each utterance - which is a feature of the West-Irish accent.

Word	IPA	West	IPA
Brain	b / r / e i / n	Bray-in	'b / r / e i / I / n
Train	t / r / e i / n	Tray-in	't / r / e i / I / n
Crane	k / r / e i / n	Cray-in	'k / r / e i / I / n
Drain	d / r / e i / n	Dray-in	'd / r / e i / I / n

Table 5: 'Irish-English' phonetical spelling by West of Ireland female.

Word	IPA	East	IPA
Leap	l / I : / p	Lee-ip	l / I : / I / p
Leave	l / I : / v	Lee-iv	l / I : / I / v
Weep	w / I : / p	Weep	w / I : / p
Sleep	s / l / I : / p	Sleep	s / l / I : / p

Table 6: 'Irish-English' phonetical spelling by East of Ireland male.

IV DISCUSSION

This short study aimed to determine whether the phonological differences within 'Irish-English' speech highlighted by Hickey, altered the pronunciation of an utterance thus resulting in making it more (or less) intelligible. The utterances that were used in the online perceptual test provided

reasonable evidence that the understanding of the word could be altered due to the regional nuances applied to them. Examples of this would be the disyllabification of the long high vowels and the enforced glottal stops in the words “*brain*” and “*leap*” as these subtle gradations provided an added forced stop within the utterance thus reducing the duration of the unvoiced vowel.

The limitations of this study would be that only two subjects were used. Future studies would require a larger group of users from various regions throughout the country in order to evaluate what effect the ‘Irish-English’ had on the intelligibility of electrolarynx speech.

REFERENCES

- [1] Kortmann, B. et al. (2004). “*Handbook of varieties of English. Volume 1: Phonology. Berlin: Mouton de Gruyter*” 68-97.
- [2] <http://www.cancer.gov/cancertopics/wyntk/larynx/page8> (as of 2/4/2012).
- [3] Houston K.M., Hillman R.E., Kobler J.B., Meltzner G.S., 1999. *Development of sound source components for new electrolarynx speech prosthesis*. IEEE Proc. International Conference on Acoustics, Speech, and Signal Processing Vol. 4, Issue, 15-19 Page(s):2347 - 2350 vol.4.
- [4] Shoureshi R.A., Chaghajerdi A., Aasted C., Meyers A., 2003. *Neural-based prosthesis for enhanced voice intelligibility in laryngectomees*. IEEE/EMBS Proc. 1st Int. conference on neural engineering, Capri Island, Italy.
- [5] Liu H., Zhao Q., Wan M., Wang S., 2006 *Enhancement of electrolarynx speech based on auditory masking*. IEEE transactions on Biomed. Eng. Vol. 53, No. 5.
- [6] Cole D., Sridharan S., Mody M., Geva S., 1997 *Application of noise reduction techniques for alaryngeal speech enhancement*. IEEE TENCON – Speech and image technologies for computing and telecommunications. 491-494.
- [7] Uemi, N., Ifukube, T., Takahashi, M., & Matsushima, J. (1994). Design of a new electrolarynx having a pitch control function, IEEE Workshop on Robot and Human, 198–202.
- [8] Ma, K., Demirel, P., Espy-Wilson, C., & MacAuslan, J. (1999). Improvement of electrolarynx speech by introducing normal excitation information. Proceedings of the European Conference on Speech Communication and Technology (EUROSPEECH), Budapest 1999, 323–326.
- [9] Gray, S., & Konrad, H. R. (1976). Laryngectomy: Postsurgical rehabilitation of communication. Archives of Physical Medicine and Rehabilitation, 57, 140–142.
- [10] Hillman, R. E., Walsh, M. J., Wolf, G. T., Fisher, S. G., & Hong, W. K. (1998). Functional outcomes following treatment for advanced laryngeal cancer. Part I–Voice preservation in advanced laryngeal cancer. Part II– Laryngectomy rehabilitation: The state of the art in the VA System. Research Speech-Language Pathologists. Department of Veterans Affairs Laryngeal Cancer Study Group. Annals of Otolaryngology and Laryngology Supplement, 172, 1–27.
- [11] Crystal, T. H., & House, A. S. (1982). Segmental durations in connected speech signals: Preliminary results. Journal of the Acoustical Society of America, 72, 705–717.
- [12] Mitchell, H. L., Hoit, J. D., & Watson, P. J. (1996). Cognitive–linguistic demands and speech breathing. Journal of Speech and Hearing Research, 39, 93–104.
- [13] Egan, J. (1948). Articulation testing methods. Laryngoscope, Vol. 58(9) , pp. 955-991
- [14] Weiss, M. S., Yeni-Komshian, G. H., & Heinz, J. M. (1979). “*Acoustical and perceptual characteristics of speech produced with an electronic artificial larynx*”. Journal of the Acoustical Society of America, 65 , 1298–1308.
- [15] Weiss, M. S., & Basili, A. G. (1985). “*Electrolaryngeal speech produced by laryngectomized subjects: Perceptual characteristics*”. Journal of Speech and Hearing Research, 28 , 294–300.
- [16] Madden, B., Nolan, M., Burke, E., Condrón, J., Coyle. (2010). “*Intelligibility of electrolarynx speech using a novel actuator*”. Proceedings from the Irish Signals and Systems conference, Cork.
- [17] Madden, B., Nolan, M., Burke, E., Condrón, J., Coyle. (2011). “*Intelligibility of electrolarynx speech using a novel hands-free actuator*”. Proceedings from the Biosignals conference, Rome.
- [18] Bliss, A.J. (1979). “*Spoken English in Ireland 1600-1740. Twenty-seven representative texts assembled and analysed*”. Dublin: Canedeus Press.