

The Differences in Prosodic Features of Alaryngeal Speech after Laryngectomy

Ljiljana Širić^{1,2*}

¹Department of Otorhinolaryngology and Head and Neck Surgery, University Hospital Centre Osijek, Osijek, Croatia

²Department of General and Applied Kinesiology, Faculty of Kinesiology, University of Zagreb, Zagreb, Croatia

*Corresponding Author: Ljiljana Širić, Department of Otorhinolaryngology, Head and Neck Surgery, Osijek University Hospital Centre, J Huttlera 4, 31000 Osijek, Croatia

Received: June 14, 2021

Published: June 25, 2021

Abstract

Purpose: Defining the prosodic features of alaryngeal speech and determining their differences depending on the type of alternative speech.

Methods: The study included 60 laryngectomized subjects of both sexes with a mean age of 63 years. The subjects were divided into three groups depending on the type of alternative alaryngeal speech. Prosodic features were assessed using a four-component scale for the assessment of prosody of alaryngeal speech by three examiners. Statistical data processing was done using the SPSS statistical package.

Results: Significant differences were found between tracheoesophageal, esophageal, and electro laryngeal speech with respect to melody and accent performance, and variable rhythm, while the existing differences in the realization of pauses in speech were not significant. Most tracheoesophageal subjects have the appropriate melody, as well as the accent achieved during the speech, the logical pauses present, and in most cases the appropriate speech rhythm.

Conclusion: By evaluating prosody, vocal rehabilitation by the installation of a tracheoesophageal prosthesis has proven to be the most optimal rehabilitation method.

Keywords: Alaryngeal Speech; Assessment; Laryngectomy; Prosody

Introduction

Changes and disorders of anatomical relationships and functional activity in the neck area affect the proper performance of multiple functions simultaneously [1]. Loss of the larynx as a voice generator has lasting effects on the laryngectomized person, and the inability to speak loudly leads to a devastating decrease in the quality of life of the patient and causes frustration with their impaired ability to communicate effectively and successfully with the environment, therefore voice rehabilitation is primary in the postoperative period. Laryngectomized persons are rehabilitated using one of the three existing methods of alternative alaryngeal voice: implantation of the tracheoesophageal prosthesis and production of the tracheoesophageal voice and speech, establishment of the esophageal voice and production of the esophageal speech, or the usage of electrolarynx and electrolaryngeal voice production. The primary purpose of rehabilitation is to enable verbal communication and to reintegrate the laryngectomized patient into the social environment [2]. Consequently, alternative speech must meet

certain qualitative criteria that, despite different environmental communication conditions, allow for adequate social contact and independent functioning of the laryngectomized person. One of these important criteria are the prosodic features of speech that affect speech intelligibility, the semantic component, and the ultimate context of the spoken [3]. Prosodic features have multiple names and divisions depending on theorists and researchers [4,5], and also, encompass multiple speech layers, from words and sentences to expressiveness and screech. The Croatian accent system is standardized as a four-pitch accent system that contains short rising, short falling, long rising and long falling pitch accents, and with them phonologically short and long unaccented syllables. Isographic heterophones are words that differ solely in accent and are produced by the opposition of all accent combinations, while words of the same segmental structure that might differ in opposition to all four accents do not exist. The place of emphasis is linguistically - rhythmically related to a particular syllable, which

means that it is impermanent and changeable in the paradigm. The accentuated length belongs to the base of the word or to a form and formation continuation. An accent unit is a prosodic word composed of one accent word or one accent and one or more non-accents [6].

Theoretically, intonation is described on the basis of phonological description, therefore different theories postulate the number of phonological entities and associate phonetic realization with them. Most intonation models' descriptions are based on the course of the movement of the tone, although there is also an intonation model's description based on the description of the goal. During subjective listening, through perception of accents, auditory differences of rising and falling pitch accents are noticed, and perception can have a phonological or acoustic description. While doing so, deviations from the norm by tone, location and duration are analyzed. During auditory quantification, the impact of sentence intonation on word prosody is questioned, because intonation modifies the accent of words in connected speech [7]. Socio-phonetic studies of the degree of desirability of accent forms investigate and observe the differences between the two accent characters by listening, determining whether the codified norm agrees with the verified one. The accents have an accent unit which is the maximum accent frame and an accent unit which is the minimum accent frame. The optimal domain size for accentuation is still not precisely defined, but in Croatian speech the most common are three-syllable and two-syllable words and given that the very first syllable is often stressing, it is possible that the dactyl and trochaic rhythmic rates are equally canonical. The optimum for accentuation as the maximum accent frame is considered to be a three-syllable word but a two-syllable word is sufficient for the analysis [8]. The degree of interaction of lexical and postlexical prosody indicates that the rising accent has a greater tendency to be preserved than falling and a greater degree of word prominence in the intonation phrase, resulting in better preservation of lexical prosody [9].

Aim of study

The aim of this study was to identify, evaluate, and compare the prosodic features of tracheoesophageal, esophageal, and electro laryngeal speech, as they play a significant role in speech intelligibility.

Subjects and Methods

The study was conducted at the Osijek University Hospital Center at the Department of Otorhinolaryngology, Head and Neck Surgery, and was approved by the Ethics Committee of the Osijek University Hospital Center. The study included 60 laryngectomized subjects of both sexes, with a median age of 63 years. The subjects were divided into three groups depending on the type of alternative alaryngeal voice and speech used daily. The first group consisted of subjects who had a prosthesis implanted, the second group consisted of esophageal-speaking subjects, and the third group

consisted of subjects using electrolarynx. All tasks scheduled for examining speech prosody were pre-written. Prosodic features were assessed using a four-component scale for the assessment of prosody of alaryngeal speech. The assessment was performed by three evaluators of the same professional competences. Four prosodic features were assessed: melody, accent, pauses, and rhythm. Each characteristic was qualitatively and temporally described according to a pre-assembled closed-type scale with grades on a Likert scale of 1 to 3. The evaluators assessed prosody during the respondent's reading of pre-prepared test material. The test material for evaluating speech prosody consisted of a series of sentences for evaluating speech intonation, accent, pauses, and rhythm of the same syntactic structure, and different into native form and arrangement of punctuation marks or pairs of sentences of different syntactic structure, and the same into native form.

Statistical methods

After the data were collected, statistical processing was performed using the SPSS computer program (version 16.0, SPSS Inc., Chicago, IL, USA). Differences of categorical variables were tested by Fisher's exact test, and differences of numerical variables between two independent groups were tested by Mann - Whitney U test with respect to deviations from the normal distribution. Differences in the abnormally distributed numerical variables of the three groups were tested with the Kruskal - Wallis test. All p values are two sided. The significance level was set at $p = 0.05$.

Results

In subjects with electrolarynx, the melody was significantly always uniform (Fisher's exact test, $P < 0.001$), not variable (Fisher's exact test, $P < 0.001$), and not appropriate (Fisher's exact test, $P < 0.001$) compared to subjects with a speech prosthesis whose melody is never monotonous. In esophageal subjects, the melody is significantly more variable, and it is significantly and always appropriate in subjects with a speech prosthesis (Table 1). The majority of subjects with electrolarynx were less likely to have an accent, 14/20 of them (Fisher's exact test, $P < 0.001$). Of the 15 (25%) subjects who only occasionally have a partial accent, there are significantly fewer esophageal subjects (Fisher's exact test, $P = 0.004$), and those with a prosthesis are more likely to (Fisher's exact test, $P < 0.001$) always have the accent (Table 2). Only 3 (5%) of subjects never had pauses, occasional respiratory breaks occurred in 19 (32%) of subjects, and 50 (83%) of subjects always had logical breaks. There were no significant differences between the groups of subjects (Table 3). The highest number of subjects, 19 (32%) had an inappropriate rhythm (slow or fast), significantly more frequently in the esophageal subjects (Fisher's exact test, $P = 0.005$). Without a variable rhythm, there were significantly more subjects with electrolarynx, 16/20 of 37 (62%) subjects (Fisher's exact test, $P = 0.001$). There are always 14/20 subjects with a speech prosthesis with an appropriate rhythm, which is significantly more than those of the esophageal or those with electrolarynx (Fisher's exact test, $P = 0.02$), (Table 4).

Table 1: Subjects according to melody and groups.

Melody	Number (%) fo Subjects				p-value
	Prosthesis	Esophageal	Electrolarynx	Total	
Monotonous					
Always	0	0	20/20	20 (33)	< 0,001
Occasionally	7/20	10/10	0	17 (28)	
Never	13/20	10/20	0	23 (38)	
Rising					
Always	4/20	3/20	0	7 (12)	< 0,001
Occasionally	4/20	11/20	1/20	16 (27)	
Never	12/20	6/20	19/20	37 (62)	
Falling					
Always	11/20	8/20	0	19 (32)	< 0,001
Occasionally	8/20	12/20	1/20	21 (35)	
Never	1/20	0	19/20	20 (33)	
Total	20/20	20/20	20/20	60 (100)	

Table 2: Subjects according to accent and groups.

Accent	Number (%) of Subjects				p-value
	Prosthesis	Esophageal	Electrolarynx	Total	
No accent					
Always	0	2/20	14/20	16 (27)	< 0,001
Occasionally	2/20	8/20	6/20	16 (27)	
Never	18/20	10/20	0	28 (47)	
Partially present accent					
Always	4/20	10/20	0	14 (23)	0,004
Occasionally	5/20	4/20	6/20	15 (25)	
Never	11/20	6/20	14/20	31 (52)	
Present accent					
Always	11/20	4/20	0	15 (25)	< 0,001
Occasionally	6/20	14/20	3/20	26 (43)	
Never	0	2/20	17/20	19 (32)	
Total	20/20	20/20	20/20	60 (100)	

Table 3: Subjects according to pauses and groups.

Pauses	Number (%) of Subjects				p- value
	Prsthesi	Esophageal	Electrolarynx	Total	
No pauses					
Always	0	0	20-Mar	3 (5)	0,10
Occasionally	20/20	20/20	17/20	57 (95)	
Respiratory pauses					
Occasionally	6/20	10/20	3/20	19 (32)	0,07
No pauses	14/20	10/20	17/20	41 (68)	
Logical pauses					
Always	18/20	15/20	17/20	50 (83)	0,57
Occasionally	2/20	5/20	3/20	10 (17)	
Total	20 (100)	20 (100)	20 (100)	60 (100)	

Table 4: Subjects according to rhythm and groups.

Rhythm	Number (%) of Subjects				p- value
	Prosthesis	Esophageal	Electrolarynx	Total	
Inappropriate (too slow or too fast)					
Always	1/20	0	6/20	7 (12)	0,005
Occasionally	4/20	11/20	4/20	19 (32)	
Never	15/20	9/20	10/20	34 (47)	
Variable					
Always	0	6/20	3/20	9 (15)	0,001
Occasionally	5/20	8/20	1/20	14 (23)	
Never	15/20	6/20	16/20	37 (62)	
Appropriate					
Always	14/20	6/20	10/20	30 (50)	0,02
Occasionally	5/20	12/20	4/20	21 (55)	
Never	1/20	2/20	6/20	9 (15)	
Total	20/20	20/20	20/20	60	

Discussion

The assessment of rehabilitation and speech of laryngectomized persons is influenced by a number of factors, and it should be considered that the experts' evaluations differ significantly from those of patients. It is certain that experts evaluate speech in a complex and structured way, unlike patients, but the quality of communication, and of life in general, is a subjective and individual category. Speech intelligibility has been the most commonly questioned, since alaryngeal speech should, in the first place, be intelligible, and secondarily, the quality of the alaryngeal voice is essential. But in practice, these components are difficult to detach because of the interplay. Jongmans' research on tracheoesophageal speakers showed that the production of initial voices and syllables was significantly more difficult than the production of final voices and syllables during rehabilitation [10]. Practice has shown that in the case of esophageal speakers, the opposite is true, whereas patients who speak with the help of electrolarynx have no such difficulty. Also, in esophageal and tracheoesophageal speech, the distinction between voiced and voiceless sounds is difficult to achieve, and the guttural fricative / h / is often omitted in spontaneous expression. Most commonly, difficulties are noted in the production of fricatives, and occasionally the intelligibility of vowels is compromised, while the production of nasals is mostly undisturbed [11]. The results of the Pols study show 72% consonant intelligibility and 74% vowel intelligibility [12]. Such results can be explained by the lack of standardized speech assessment scales in this population or existing limited and non-standardized scales that evaluate only certain aspects of speech. McColl states that changes after laryngectomy affect the change in the position of the tongue and the length of the vocal tract, which

further affects the accuracy of pronunciation. It often happens that alaryngeal speakers produce only certain phonemes, thereby compromising articulation. Consequences on acoustic features are seen as the reduction or deviation of acoustic - phonemic features that the listener relies on during the conversation [13]. Roozen primarily investigated the intelligibility of tracheoesophageal speech in spontaneous speech situations, but in the discussion referred to prosodic features, noting that there were significant discrepancies in speech intonation and duration of spontaneous verbal expression [14]. According to the prosodic features, the results of this study show significant differences between tracheoesophageal, esophageal and electro-laryngeal speech with respect to melody and accent realization, and a variable rhythm, while existing differences in the realization of pauses in speech are not significant. The majority of tracheoesophageal subjects have the appropriate melody, as well as the accents that have always been achieved during speech, the always-occurring logical pauses, and the most commonly appropriate speech rhythm, which is in favor of surgical rehabilitation by the installation of a tracheoesophageal prosthesis. The least achieved prosodic features according to the assessment of melody and accent have the subjects with electrolarynx, and according to the evaluation of pauses and rhythm the esophageal subjects. Occasionally, a number of subjects have respiratory breaks, regardless of the type of alternative speech and no significant difference between the groups. Such a result is not expected with tracheoesophageal speech due to the fact that they use air from the lungs for speech, but it can be explained by the increased arousal of the subjects during participation in the study or the pursuit of a faster speech rhythm where the speaker, due to the lack of time, does not take the speech air where they logically should, but where they are forced due to the lack of air.

Conclusion

According to the examined criteria, vocal and speech rehabilitation with the installation of a tracheoesophageal prosthesis proved to be the most optimal method of rehabilitation. The production of tracheoesophageal voice and speech registered the most optimal realization of prosodic features, which significantly affects the level of intelligibility of spontaneous expression. Consequently, it is logical that implantation of the prosthesis after laryngectomy is the primary and best choice of rehabilitation, but it should not be forgotten that the other two methods have their potential beneficiaries in cases where this method is not an option.

References

1. Prgomet D (2020) Tumori glave i vrata. Medicinska naklada, Zagreb p. 201-236.
2. Širić Lj (2020) Rehabilitacija glasa i govora nakon operacije grla. Bajtl V, Pavičić Lj (Eds.), Vodič kroz Karcinom Grkljana Osijek, KLOOS, KBCO, Gradska liga protiv raka p. 35-45.
3. Širić Lj, Rosso M, Včev A (2018) The Role of Esophagus in Voice Rehabilitation of Laryngectomees. Chai J (Eds.), Esophageal Cancer and Beyond. Intechopen, London, UK p. 67-82.
4. Škarić I (2001) Razlikovna prozodija. Jezik 48(1): 11-19.
5. Möbius B (1993) Ein quantitatives Modell der deutschen Intonation. Analyse und Synthese von Grundfrequenzverläufen. Max Niemeyer Verlag, Tübingen.
6. Pletikos E (2008) Akustički opis hrvatske prozodije riječi. Sveučilište u Zagrebu: Filozofski fakultet.
7. Pletikos E (2008) Akustičke i perceptivne osobine naglasaka riječi u hrvatskim nad dijalektalnim govorima. Tošović B (Eds.). Die Unterschiede zwischen dem Bosnischen/Bosniakischen, Kroatischen und Serbischen, Slavische Sprachkorrelationen. Lit Verlag, Wien - Berlin p. 404-429.
8. Škarić I (2006) Hrvatski govorili! Zagreb, Školska knjiga.
9. Josipović V (1995) Akcenatska prozodija i dvotonski pristup intonaciji. Suvremena lingvistika 21(40): 51-79.
10. Jongmans P (2008) The intelligibility of tracheoesophageal speech: an analytic and rehabilitation study. Print Partners Ipskamp, Enschede.
11. Batstone MD, Scott B, Lowe D, Rogers SN (2009) Marginal mandibular nerve injury during neck dissection and its impact on patient perception of appearance. Head Neck 31: 673-678.
12. Pols LCW (1983) Three-mode principal components analysis of confusion matrices, based on identification of Dutch consonants, under various conditions of noise and reverberation. Speech Commun 2: 275-293.
13. McColl DA (2006) Intelligibility of tracheoesophageal speech in noise. J Voice 20(4): 605- 615.
14. Roozen M (2005) The intelligibility of tracheoesophageal speech in spontaneous speech situations. Universiteit van Amsterdam, Netherlands.



This work is licensed under Creative Commons Attribution 4.0 License

To Submit Your Article Click Here: [Submit Article](#)

DOI: [10.32474/SJO.2021.06.000250](https://doi.org/10.32474/SJO.2021.06.000250)



Scholarly Journal of Otolaryngology

Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles