

■ Listeners' Social Perception of Speakers after Treatment for Laryngeal Cancer

■ La perception sociale par des auditeurs, de locuteurs ayant reçu un traitement contre un cancer du larynx

Aarthi Turcotte
 Anroup Wilson
 Jeffrey Harris
 Hadi Seikaly
 Jana M. Rieger

Aarthi Turcotte, MSc,
 Calgary Health Region,
 Calgary, Alberta

Anroup Wilson, MSc,
 Department of
 Communication Disorders,
 Deer Lodge Centre,
 Winnipeg, Manitoba

Jeffrey Harris, MD, FRCSC,
 Division of Otolaryngology
 – Head and Neck Surgery,
 University of Alberta,
 Edmonton, Alberta
 Faculty of Medicine and
 Dentistry, University of
 Alberta, Edmonton, Alberta

Hadi Seikaly, MD, FRCSC
 Division of Otolaryngology
 – Head and Neck Surgery,
 University of Alberta,
 Edmonton, Alberta
 Faculty of Medicine and
 Dentistry, University of
 Alberta, Edmonton, Alberta

Jana M. Rieger, PhD
 Institute for Reconstructive
 Sciences in Medicine
 (iRSM), Edmonton,
 Alberta
 Faculty of Rehabilitation
 Medicine, University of
 Alberta, Edmonton, Alberta

Abstract

Three primary methods to treat laryngeal cancer include radiation therapy, total laryngectomy, and supracricoid laryngectomy. Perceptual assessment to determine the social impact of vocal outcomes related to each treatment was employed to understand the effect that a disordered voice may have on societal perception of patients undergoing these treatments. Forty listeners were recruited to rate the voices of four groups of speakers: individuals treated with radiation therapy; individuals treated with total laryngectomy and rehabilitated with a tracheoesophageal voice prosthesis; individuals treated with supracricoid laryngectomy; and individuals with no history of vocal disorders. The listeners rated the speakers' voices on rating scales with eight descriptors related to the social perception of speech. The ratings were made on 7-point Likert scales. Results indicated that listeners rated the non-surgical groups more positively than the surgical groups on traits such as *attractive*, *clever*, *sophisticated*, and *trustworthy*. The surgical groups were rated more negatively than non-surgical groups on traits such as *scary*, *annoying*, and *intimidating*. These findings provide insight into how patients undergoing these treatments may be perceived by society. Individuals with total laryngectomies may be at a higher risk of encountering negative social stigmatization in their daily life.

Abrégé

Les trois principales méthodes pour traiter un cancer du larynx comprennent la radiothérapie, une laryngectomie totale et une laryngectomie supracricoïdienne. On a mené une évaluation subjective de l'incidence sociale de la voix résultant de chaque traitement afin de comprendre l'effet d'un trouble de voix sur la perception qu'a la société des patients qui subissent ces traitements. On a recruté 40 auditeurs pour évaluer la voix de quatre groupes de locuteurs : les personnes ayant subi une radiothérapie, celles ayant eu une laryngectomie totale et une réadaptation avec un implant pour la voix trachéo-oesophagienne, celles ayant subi une laryngectomie supracricoïdienne et celles n'ayant pas d'antécédents de trouble de la voix. Les auditeurs ont évalué la voix des locuteurs sur des échelles à huit descripteurs portant sur la perception sociale de la voix. Ils ont utilisé une échelle de Likert à sept niveaux. Les résultats indiquent que les auditeurs ont accordé une évaluation plus positive au groupe de locuteurs n'ayant pas subi de chirurgie qu'aux groupes ayant eu une chirurgie pour les traits de personnalité comme le caractère attrayant, l'intelligence, la subtilité et la fiabilité. Les groupes ayant subi une opération ont été jugés comme faisant plus peur, comme étant plus contrariants et comme étant plus intimidants que les autres groupes. Ces résultats donnent un aperçu de la façon dont les patients qui subissent ces traitements sont perçus par la société. Les laryngectomisés totaux risquent davantage d'être l'objet de stigmatisation dans leur quotidien.

Keywords: laryngeal cancer, radiation therapy, total laryngectomy, supracricoid laryngectomy, perception, social attributes

Background

Approximately 1,100 cases of laryngeal cancer are diagnosed annually in Canada (Canadian Cancer Society/National Cancer Institute of Canada, 2008). Three methods of interest to this study used for primary treatment of laryngeal cancer include total laryngectomy, supracricoid laryngectomy, and radiation therapy. The treatment of choice is related largely to the stage of cancer at diagnosis. Each method of treatment has the tendency to distort a speaker's voice because anatomical structures necessary to produce voice may be removed or altered. Modifications to the vocal structures have the potential to change the vocal quality of the speaker, which has an impact on their quality of life (Starmer, Tippett, & Webster, 2008; Lallh & Rochet, 2000; Witt et al., 1997) and professional options (Hurst & Cooper, 1983).

Total laryngectomy becomes necessary in advanced stages of the cancer (Zacharek et al., 2001). This surgery involves the removal of the entire larynx (including the vocal folds). Speech rehabilitation must then focus on replacing the voice, which is the source signal for speech. This can be achieved with either electrolaryngeal, esophageal, or tracheoesophageal speech. Electrolaryngeal speech requires an electrolarynx, which is a mechanical handheld device with a vibrating plastic diaphragm. This vibrating diaphragm is placed against the neck and the sound waves are transmitted through the neck tissue into the vocal tract. Using this artificial voicing source, the patient can then move the articulators and produce speech sounds. A second choice for speech production is esophageal speech. The patient must insufflate air into the upper esophagus by injecting, swallowing, or "inhaling" air just below the level of the cricopharyngeal segment. When the air is released again, the upper esophageal sphincter vibrates and a ructus sound is produced. This ructus can be used as the sound source for speech. The third alternative is tracheoesophageal (TE) speech. In order to enable a laryngectomee to use TE speech, a surgically created fistula connects the trachea and esophagus (Blom & Singer, 1995). A small one-way valve, called a voice prosthesis, is placed in the fistula. The prosthesis prevents material from the esophagus from spilling into the trachea while allowing air to flow from the trachea into the esophagus. TE speech can then be produced by using air from the lungs to drive the pharyngoesophageal segment. In order to speak, the stoma must be sealed (with the thumb or with a one-way valve) during exhalation to divert the air into the esophagus.

One more recent form of treatment for laryngeal carcinoma is the supracricoid laryngectomy (SCL; Farrag et al., 2007). This treatment entails the removal of a portion of the larynx above the cricoid cartilage. At least one arytenoid cartilage must remain intact to allow for voluntary ad- and abduction of the vocal folds. Typically, the remaining arytenoid cartilage/s will move in an

anterior direction to make contact with the epiglottis or the base of the tongue in order to produce voice. Speech is produced without a mechanical aid or a stoma in the neck (Schindler et al., 2005; Coman, Grigg, Tomkinson, & Gallagher, 1998).

Primary radiation therapy is used for smaller (early stage) laryngeal tumours. This particular approach is a non-invasive procedure that keeps the vocal tract structures relatively intact. The drawbacks to this treatment approach are the length of treatment period and the unavoidable irradiation of healthy cells during treatment (Fietkau & Sauer, 1992).

The voices of individuals who have undergone treatment for laryngeal cancer are often characterized as abnormal. One study showed that patients using TE speech were judged to have a less acceptable voice quality than normal control speakers (Pindzola, Auburn, & Cain, 1988). Finizia, Dotevall, Lundström, and Lindström (1999) indicated that patients treated with radiation therapy for laryngeal cancer were rated more positively by inexperienced listeners in terms of voice quality, speech intelligibility, and speech acceptability than patients with TE speech. They also found that the laryngectomy group rated themselves lower on the same characteristics when compared to the normal controls and the radiation therapy group. Zacharek et al. (2001) illustrated that patients with SCL are judged to be highly intelligible, however, their vocal quality was found to have different dysphonic features. On a self-rating questionnaire, these patients identified themselves as having moderate-severe voice problems. Eksteen, Rieger, Nesbitt, and Seikaly (2003) conducted a study to examine the acoustic information related to individuals who had undergone treatment for laryngeal cancer, including those who had undergone either total laryngectomy, supracricoid laryngectomy, or radiation therapy. They found that the speakers who had undergone either a supracricoid laryngectomy or a total laryngectomy had higher noise-to-harmonics ratios, as well as higher jitter and shimmer values, in comparison to a group of patients who had undergone radiation therapy and a control group of normal speakers.

While it is self-evident that modifications of the vocal structures during treatment of laryngeal cancer have the potential to change the voice (Starmer et al., 2008), there is little reliable information that can be provided to patients to counsel them about the perceptual consequences of the treatment they will receive (Pindzola et al., 1988). As an altered voice quality can affect a patient's quality of life, information about the effects of the patient's chosen treatment on the voice should be provided. Different aspects of life that have been shown to be affected by a voice disorder are self-esteem, career options, and social interactions.

Research shows that when vocal characteristics deviate from the norm, there may be an impact on how speakers perceive themselves and their quality of life (Lallh & Rochet, 2000). Witt et al. (1997) investigated how children with cleft palate viewed their speech and how it affected their lives. Almost half of the children in the cleft palate group felt

that they had problems with their speech and expressed a desire to improve their speech.

Disorders of speech may also influence occupational prospects. Hurst and Cooper (1983) looked at how stuttering affected future professional options. They sent out questionnaires to employers and asked their opinions about job performance and employability for individuals who stutter. While the majority of employers rejected the idea that stuttering would interfere with job performance, they felt that stuttering could affect employability and promotion prospects. Similarly, individuals with voice disorders have expressed the opinion that their voice had impinged on their occupational prospects (Smith et al., 1996).

Listeners often stigmatize speakers with speech disorders by associating their speech with negative characteristics. For example, Witt et al. (1997) found that parents and teachers consistently rated children with cleft palate more negatively than non-cleft children. In another study, McKinnon, Hess, and Landry (1986) investigated how college students reacted to moderately disordered speech samples of stuttering, hypernasality, and lateral lisp compared to normal speech. Listeners felt higher anxiety and preferred to disassociate and distance themselves from speakers with speech disorders. Blood, Mahan, and Hyman (1979) also found that listeners were reluctant to communicate with individuals with disordered speech. Elementary school-aged children were found to respond negatively to recordings of speakers with severe hypernasality. Lallh and Rochet (2000) found that speakers with voice disorders and hypernasal vocal qualities were rated more negatively than speakers with normal voices on characteristics such as intelligence, reliability, kindness, and physical appearance. Finally, a study by Rieger et al. (2006) documented listeners' social perceptions of speakers both before and after surgery for oropharyngeal cancer and found that positive perceptions of speakers were significantly diminished after surgery, whereas negative perceptions were increased.

When carrying out social perceptual studies, it is important that the demographics and characteristics of the listeners are rigorously controlled. Research has indicated that age and sex may impact how listeners perceive vocal quality. Williams and Dietrich (2001) found that as the age of listeners decreased, their ratings of individuals with communication disorders became more positive. In another study, Deal and Oyer (1991) found that female listeners tended to rate speakers who stutter more positively than male listeners. Thus, care must be taken when interpreting studies that have not taken listener characteristics into account.

Present Study

While acoustic outcomes and intelligibility of speech have been compared across the three methods of laryngeal cancer treatment that are of interest to this study (Eksteen et al., 2003), studies to date have not compared the social perception of the vocal outcome. The importance of understanding how individuals will be regarded in society after treatment for laryngeal cancer cannot be overemphasized.

Table 1
Listener Characteristics

Listener Number	Occupation	Sex	Age
1	Professional	Male	23
2	Professional	Male	21
3	Professional	Male	23
4	Student	Male	18
5	Professional	Male	26
6	Professional	Female	29
7	Other	Female	46
8	Student	Female	23
9	Student	Female	20
10	Student	Female	22
11	Professional	Male	24
12	Other	Male	22
13	Professional	Female	23
14	Student	Female	20
15	Student	Male	23
16	Professional	Male	30
17	Student	Female	18
18	Student	Female	22
19	Student	Male	26
20	Student	Female	21
21	Student	Female	19
22	Student	Female	24
23	Student	Male	23
24	Student	Female	18
25	Professional	Male	25
26	Student	Female	18
27	Student	Female	22
28	Other	Male	23
29	Student	Female	20
30	Student	Male	26
31	Professional	Female	23
32	Student	Male	24
33	Professional	Female	22
34	Student	Male	23
35	Professional	Female	23
36	Professional	Male	22
37	Professional	Male	56
38	Other	Female	57
39	Student	Male	23
40	Professional	Male	50

It has been observed that while speech intelligibility may remain within normal limits, the produced speech may not be perceived as socially acceptable. These perceptions may impact the patients' social relationships as well as their occupational options. The purpose of this study was to assess listeners' social perceptions of patients following three different types of treatment for laryngeal cancer. Previous studies have suggested that perceptual rating scales can realistically reflect how listeners judge voices (van As, Koopmans-van Beinum, Pols, & Hilgers, 2003; Qi & Weinberg, 1995). Therefore, the main tool for measurement in this study is a rating scale that was devised to assess how patients will be viewed by the general population following treatment for laryngeal cancer.

There were three specific research questions for this study. Firstly, the study aimed to determine if listeners would rate the speakers who had undergone treatment for laryngeal cancer differently than a control group. Secondly, it was of interest to determine whether one type of treatment for laryngeal cancer resulted in better social perceptions. Thirdly, it was of interest to determine if the demographics (i.e., sex, occupation, and age) of the listeners influenced the social perception ratings.

Methods

Participants

Forty adult participants, 20 males and 20 females, were recruited as listeners for this study. The average age of the participants was 26 years and ranged from 18-57 years (Table 1). The participants came from student, professional, and other backgrounds. The category *students* was composed of currently-enrolled university students, the category *professional* consisted of individuals who were working in a full-time capacity outside the home, and the category *other* consisted of adults who were not currently employed outside the home. Participants were all English-speaking and residing in Canada. None of the listeners were aware of the nature of the communication disorder or the medical history of the speakers to which they would be listening. Participants were informed that they were taking part in a research study where they would be required to make judgments about social perceptions of speech.

Methods

Speech Recordings: Voice samples from 22 speakers used in this study came from a database at the Institute for Reconstructive Sciences in Medicine (iRSM) at the Misericordia Community Hospital in Edmonton, Alberta. The voice samples were used in a previous study (Eksteen et al., 2003) approved by the Health Review Ethics Board at the University of Alberta. Ethics clearance to use the archival data for the current study was obtained. All voice samples used in this study came from male speakers. They were classified as belonging to one of four groups: six speakers were in the total laryngectomy (TL) group, six were in the supracricoid laryngectomy (SCL) group,

Table 2

Speaker Characteristics

Speakers	Age	Treatment Type
1	45	SCL
2	57	SCL
3	59	SCL
4	63	SCL
5	69	SCL
6	75	SCL
7	48	TL
8	58	TL
9	62	TL
10	63	TL
11	68	TL
12	69	TL
13	55	RT
14	64	RT
15	69	RT
16	81	RT
17	83	RT
18	81	Control
19	73	Control
20	56	Control
21	72	Control
22	49	Control

Notes: SCL = supracricoid laryngectomy; TL = total laryngectomy; RT = radiation therapy

five were in the radiation therapy (RT) group, and five speakers without a history of laryngeal pathology formed a control group (Table 2). The TL speakers used a tracheoesophageal voice.

Stimulus CD: A CD was produced by transferring the speakers' voice recordings from a digital audiotape (DAT) onto a computer via the Computerized Speech Lab (CSL model 4400, Kay Elemetrics, Pine Brook, NJ). The speech recordings were adjusted for loudness across all speech samples. Both extraneous pauses as well as extended pauses were removed from the speech samples. The voice samples were randomly ordered and recorded onto the CD. There were 26 samples. Of these, four were repeated to obtain intra-rater reliability estimates. The repeated samples were comprised of two speakers from the control group, one speaker from the RT group and one speaker from the SCL group.

Two speech samples from speakers with no history of vocal pathology were included for listener practice. The practice speakers read a sentence from the Rainbow

Passage ("These take the shape of a long round arch with its path high above and its two ends apparently beyond the horizon"). The intent of the practice samples was to familiarize the listeners with the content of the speech sample and the content of the rating form. In addition, it gave the listeners a chance to adjust the volume of the CD to a comfortable listening level.

An announcement informing the participants that they could adjust the volume levels to ensure a comfortable listening level during the study was played. Participants were instructed to verify that the speaker number at the top left corner of the rating scale page matched the spoken number preceding each voice sample on the CD. Instructions about filling out the scale were given to the participants via an announcer who stated, "Listen to each voice and circle the number for the characteristic that best corresponds to the voice that you have just heard." When the CD continued, the announcer said, "Sample speaker number one," which was followed with a two second pause. The first practice speaker then read the selected sentence from the Rainbow Passage, which was followed by a 35 second pause to allow the listeners to complete their rating judgments on the social perception scale provided. The same procedure was repeated for the second practice speaker which concluded the two practice trials. Once the practice trials were completed, the announcer stated, "Now we are ready to begin the study." The same procedure that was utilized for the practice samples was also used for the 22 experimental speaker samples and for the four repeated samples that were included for intra-rater reliability. At the end of the CD, the announcer thanked the listeners for their participation in the study.

Social Perception Scale: The scale was produced to evaluate listeners' impressions of individuals that have undergone one of three laryngeal cancer treatments. The scale contained eight adjectives, four that were positive qualities and four that were negative. The selected adjectives were: trustworthy, attractive, clever, sophisticated, boring, scary, intimidating, and annoying (Lallh & Rochet, 2000; Bloom, Zajac, & Titus, 1999; Blood et al., 1979). For every adjective, there was an associated 7-point scale where the anchors were 1 (*not at all*) and 7 (*very much*).

Portable CD players (Sony D-EJ621) were used to play the CD recordings to listeners. The CD players had a frequency response of 20–20 000 Hz. To allow for effective listening, participants used headphones (Jensen JF25) that had a frequency response matching that of the CD players (20–20 000 Hz).

Participants were either tested alone or in pairs. When tested in pairs, the listeners were seated so that they could not see each other. A hearing screening was conducted prior to the listening task. If any participant was found to have a hearing loss, they were excused from the study and were referred to an audiologist for further testing. After the hearing screening, other demographic information was collected such as the participants' age, sex, and occupation.

The participants were given a few minutes to become familiar with the adjectives that were used in the social perception scale. Following this, the headphones were placed on the listener's head and adjusted for comfort. The experimenter then pressed play to begin the study. The experimenter remained in the room for the full duration of the CD playing to ensure that there were no problems with the equipment and to guarantee that the CD was played without being stopped.

Statistical Analysis

SPSS was used to perform all statistical calculations. Forty listeners rated 26 voice files on each of the eight descriptive variables. For each descriptive variable, the ratings from the 40 listeners were averaged and calculated for each of the 22 speakers. The listeners' second ratings for the four randomly selected speakers were removed from the main analysis and used only for estimating intra-rater reliability. A one-way multivariate analyses of covariance (MANCOVA) was employed to determine if significant differences between listener ratings of the speech samples existed. This allowed for between-group comparisons to be made, while controlling for age of the speakers. The dependent variables were the social perception categories (boring, clever, etc.), the between-group factor was treatment group, and the covariate was age.

A two-way MANCOVA was conducted to determine the effect of listener characteristics on their ratings of the speakers. The two factors were sex (male or female) and occupation (student, professional, and other), and the covariate was age of the listeners.

An intra-class correlation coefficient (ICC) was completed to establish intra-rater reliability. This was done through comparison of listener ratings of four repeated speech samples, totaling 18% of the data. ICCs were also conducted to determine inter-rater reliability across listeners.

When more than two levels within the independent variable being examined existed and significant main effects were present, post hoc testing was used to determine significant group differences. This study used SPSS Bonferroni adjusted p-values where $p < .05$ for the post hoc statistics.

Results

Inter-rater Reliability

The mean inter-rater reliability across the eight variables revealed an ICC value of 0.9626 (range = 0.8904 and 0.9867; Table 3). Mean intra-rater reliability across the eight variables revealed an ICC value of 0.7335 (range = 0.497 and 0.8137; Table 4).

Effects of listener characteristics

The effect of listeners' age, sex, and occupation on perceptual ratings were examined through a two-way MANCOVA. The results indicated that there were no significant effects of these variables on how speakers were perceived.

Table 3
Inter-rater Reliability

Adjective	ICC Average Rater Value	95% Confidence Interval	
		Lower Bound	Upper Bound
Boring	0.8904	0.813	0.9466
Attractive	0.9767	0.9602	0.9886
Clever	0.9735	0.9547	0.9871
Scary	0.9867	0.9773	0.9935
Annoying	0.9586	0.9294	0.9798
Sophisticated	0.9804	0.9665	0.9904
Intimidating	0.9650	0.9402	0.9829
Trustworthy	0.9701	0.9489	0.9854

Table 4
Intra-rater Reliability

Variable	ICC Average Rater Value	95% Confidence Interval	
		Lower Bound	Upper Bound
Boring	0.4937	0.3083	0.6294
Attractive	0.8137	0.7455	0.8636
Clever	0.8048	0.7333	0.8571
Scary	0.7295	0.6304	0.802
Annoying	0.6218	0.4833	0.7231
Sophisticated	0.8134	0.7448	0.8636
Intimidating	0.7660	0.6802	0.8287
Trustworthy	0.7624	0.6719	0.8275

Listener judgments of speakers' characteristics

Means and standard deviations of listener responses for each perceptual characteristic are reported in Table 5. Table 6 includes significant effects found between particular treatment groups through post hoc analyses, and Table 7 provides a summary of the between-group results.

Boring

A group effect was found ($p = .05$) for the variable *boring*. With respect to group means, the control group was rated to be the most boring ($M = 4.49$), followed by the RT group ($M = 3.75$), the SCL group ($M = 3.65$), and the TL group ($M = 3.39$). Post hoc analyses revealed that significant between-group differences existed only between the TL group and RT group ($p = .05$). No age effects were found for this adjective, indicating that speakers' age did not influence listener ratings.

Attractive

A group effect was found ($p < .001$) for the variable *attractive*. The TL group was rated as the least attractive ($M = 1.36$), followed by the SCL group ($M = 1.85$), the RT group ($M = 2.85$), and the control group ($M = 3.01$). Specifically, significant differences were found through post hoc analyses between the SCL group and the RT group ($p < .005$), the SCL group and the control group ($p < .005$), the TL group and the RT group ($p < .001$), as well as the TL group and the control group ($p < .001$). An age effect was found for this variable ($p = .001$),

demonstrating that speakers' ages significantly affected listeners' ratings. The older speakers were rated as less attractive than the younger speakers.

Clever

A significant group effect was found ($p < .001$) for the variable *clever*. The control group was rated the most clever ($M = 3.94$), while TL group was rated the least clever ($M = 2.28$). The perceptual ratings for the two remaining groups fell in between, with the SCL group receiving a lower rating ($M = 2.78$) than the RT group ($M = 3.68$). Post hoc analyses revealed that significant differences existed between the SCL group and the RT group ($p < .05$), the SCL group and the control group ($p = .05$), the TL group and RT group ($p < .005$), and the TL group and the control group ($p = .001$). A significant age effect was found for this adjective ($p = .01$). The older speakers were rated as less clever than the younger speakers.

Scary

A significant group effect was found ($p < .001$) for the variable *scary*. The TL group was rated as the most scary ($M = 4.69$), followed by the SCL group ($M = 4.08$), the RT group ($M = 1.97$), and the control group ($M = 1.36$). Post hoc analyses revealed significant differences between the SCL group and the RT group ($p = .001$), the SCL group and the control group ($p < .001$), the TL group and the RT group ($p < .001$), and the TL group and the control group ($p < .001$). No significant age effects were found for this variable.

Annoying

A significant group effect was revealed ($p < .001$) for the variable *annoying*. The RT group was rated the least annoying ($M = 3.00$), while the TL group was rated as the most annoying ($M = 4.81$). The control group and SCL group ratings fell in between the previously mentioned groups, with SCL group receiving a mean rating of 4.31 and the control group receiving a mean rating of 3.06. Post hoc testing revealed significant differences between the SCL group and the RT group ($p < .05$), the SCL group and the control group ($p = .05$), the TL group and the RT group ($p = .001$), as well as between the TL group and the control group ($p < .005$). A significant age effect was found for this particular adjective ($p = .05$). The older speakers were rated as more annoying than the younger speakers.

Sophisticated

A significant group effect was revealed ($p < .001$) for the variable *sophisticated*. The control group was found to be the most sophisticated ($M = 4.03$), followed by the RT group ($M = 3.47$), the SCL group ($M = 2.46$), and the TL group ($M = 2.00$). Differences were found between the following groups through post hoc testing: the SCL group and the RT group ($p < .01$), the SCL group and the control group ($p = .001$), the TL group and the RT group ($p < .001$), and the TL group and the control group ($p < .001$). A significant age effect was found for this variable ($p = .001$). The older speakers were rated as less sophisticated than the younger speakers.

Intimidating

There was a significant group effect ($p < .001$) found for the variable *intimidating*. The TL group was rated the most intimidating ($M = 3.99$), followed by the SCL group ($M = 3.57$), the RT group ($M = 2.10$), and the control group ($M = 1.77$). Post hoc measures showed that significant differences existed between the SCL group and the RT group ($p < .005$), the SCL group and the control group ($p < .001$), the TL group and the RT group ($p < .001$), and the TL group and the control group ($p < .001$). No significant age effects were found for this adjective.

Trustworthy

A significant group effect was discovered ($p < .001$) for the variable *trustworthy*. The RT group was found to be the most trustworthy ($M = 4.31$), while the TL group was rated the least trustworthy ($M = 2.43$). Accordingly, the SCL group and the control group average ratings fell between the previously mentioned groups with the SCL group receiving a mean rating of 2.75 and the control group receiving a mean rating of 4.27. Post hoc test measures indicated that the following groups had significant differences between each other: the SCL group and the RT group ($p < .001$), the SCL group and the control group ($p = .001$), the TL group and the RT group ($p < .001$), and the TL group and the control group ($p < 0.001$). No significant age effects were found for this variable.

Discussion

The purpose of this study was to determine how individuals who had undergone three different types of treatment for laryngeal cancer were perceived by naïve listeners. The voices of these individuals were compared to controls who had no history of laryngeal pathology. The naïve listeners rated all voices using a 7-point rating scale with eight descriptor adjectives, which were: *boring*, *attractive*, *clever*, *scary*, *annoying*, *sophisticated*, *intimidating*, and *trustworthy*. Significant between-group effects were found for all adjectives. In addition, some ratings appeared to be influenced by the age of the speakers. Finally, listener characteristics such as age, sex, and occupation did not appear to influence the results.

Table 5

Means and Standard Deviations

Dependent Variable	Group	Mean	Standard Deviation
Boring	SPCL	3.65	0.36
	TL	3.39	0.31
	RT	3.75	0.88
	Control	4.49	0.70
Attractive	SPCL	1.85	0.70
	TL	1.36	0.14
	RT	2.85	1.09
	Control	3.01	0.60
Clever	SPCL	2.78	0.58
	TL	2.28	0.21
	RT	3.68	1.14
	Control	3.94	0.75
Scary	SPCL	4.08	1.21
	TL	4.69	0.80
	RT	1.97	0.74
	Control	1.36	0.18
Annoying	SPCL	4.31	0.71
	TL	4.81	0.38
	RT	3.00	1.04
	Control	3.06	0.83
Sophisticated	SPCL	2.46	0.66
	TL	2.00	0.20
	RT	3.47	1.14
	Control	4.03	1.10
Intimidating	SPCL	3.57	0.59
	TL	3.99	0.77
	RT	2.10	0.46
	Control	1.77	0.41
Trustworthy	SPCL	2.75	0.55
	TL	2.43	0.24
	RT	4.31	0.87
	Control	4.27	0.37

Note: The scale was 1 = *not at all* and 7 = *very much*

Group effects

Significant effects between each treatment group were found for each dependent variable. Common trends were observed between all of the variables except for the adjective *boring*. In answering the first question of this study, whether listeners perceived differences between the treatment groups and the control group, the results generally suggest that the control group was rated more positively than the surgically treated groups, but not the radiation therapy group. In consideration of the second question, whether there were any differences between the treatment groups, significant differences were not found between individuals who had received SCL and those who received TL procedures. This indicates that listeners did not perceive a large difference between the speakers of these two surgical groups on social perceptual attributes. However, there were consistent significant differences in social perception ratings

Table 6

Post Hoc Analysis of between group differences that were significant

Characteristic	Treatment Group 1	Treatment Group 2	Significance
Trustworthy	Supracricoid	Radiation	0.000
	Supracricoid	Normal	0.001
	Laryngectomy	Radiation	0.000
	Laryngectomy	Normal	0.000
Intimidating	Supracricoid	Radiation	0.004
	Supracricoid	Normal	0.001
	Laryngectomy	Radiation	0.000
	Laryngectomy	Normal	0.000
Sophisticated	Supracricoid	Radiation	0.006
	Supracricoid	Normal	0.001
	Laryngectomy	Radiation	0.000
	Laryngectomy	Normal	0.000
Annoying	Supracricoid	Radiation	0.013
	Supracricoid	Normal	0.025
	Laryngectomy	Radiation	0.001
	Laryngectomy	Normal	0.002
Scary	Supracricoid	Radiation	0.001
	Supracricoid	Normal	0.000
	Laryngectomy	Radiation	0.000
	Laryngectomy	Normal	0.000
Clever	Supracricoid	Radiation	0.030
	Supracricoid	Normal	0.012
	Laryngectomy	Radiation	0.002
	Laryngectomy	Normal	0.001
Attractive	Supracricoid	Radiation	0.002
	Supracricoid	Normal	0.002
	Laryngectomy	Radiation	0.000
	Laryngectomy	Normal	0.000
Boring	Laryngectomy	Normal	0.054

Table 7

Summary of significant between-group results

Attribute		TL	SCL	RT
Boring	TL			
	SCL			
	RT	*		
	Control			
Attractive	TL			
Clever	TL			
Scary	SCL			
Annoying	SCL			
Sophisticated	RT	*	*	
Intimidating	RT			
Trustworthy	Control	*	*	

between the surgically-treated groups and the radiation therapy group, with the latter being rated more positively. The results suggest that listeners tended to rate individuals in the non-surgical groups more positively than individuals in the surgical groups. Thus, the voicing changes that accompany the loss of vocal fold tissue and the development of a new voicing source are picked up by the listeners and such voice changes may influence the social perception of these patients in a negative manner.

The variable *boring*, which was chosen as an adjective in an attempt to assess perception of the monotony of the speaking pattern, was rated more inconsistently by the listeners when compared to the other adjectives. A significant difference was found only between the TL group and the control group. In addition, the listener ratings appeared to result in a more positive perception of the voices of individuals with TL (i.e., they were rated as the least boring speakers out of all four groups). Initially, this appeared counterintuitive given the fact that all the speakers with TL spoke with a trachesophageal voice prosthesis and therefore used their pharyngoesophageal segment as the voicing source. This will usually result in a more monotonous voice, and thus would have been thought to result in higher ratings of *boring* than the other groups. As the listeners were not provided with a standardized definition, their own perception of the meaning of the word *boring* may have differed from one another. From the researchers' point of view, the word *boring* was initially intended to reflect monotony of a voicing source (i.e., a lack of pitch inflection). However, from an untrained listener's point of view, the voices of the TL group may have been the most interesting because they were also the most

deviant from normal, as has been established through an acoustic study of the same voices (Eksteen et al., 2003). Upon hearing the tracheoesophageal voices, the listeners may simply have reacted to the novelty of the sound.

Speaker's age effects

Significant age effects were found for the adjectives *attractive*, *clever*, *annoying*, and *sophisticated*. The listeners rated the older speakers less *attractive*, less *clever*, more *annoying*, and less *sophisticated* than the younger speakers. van As, Hilgers, Verdonck-de Leeuw and Koopmans-van Beinum (1998) analyzed voices of patients with total laryngectomies and found that younger speakers had significantly lower jitter values than older speakers. However, there were no significant differences in perceptual ratings (e.g., abnormal-normal; ugly-beautiful, etc.). Linville (1996) indicated that older speakers can exhibit increased harsh, hoarse, strained, and breathy vocal qualities. This finding may explain why the listeners in this current study may have rated older speakers more negatively on some of the descriptors.

Listener effects

Listeners' age, sex, and occupation were not found to influence how listeners rated the speakers on the listed characteristics. These findings are somewhat inconsistent with past research. Williams and Dietrich (2001) found that younger participants tended to rate individuals with voice disorders more favourably than older participants. The age range of the listeners in their study was quite similar to the age range in the current study. One potential reason for this difference may be that the participants in the Williams and Dietrich study were provided with a description of an individual with a communicative disorder, rather than a voice sample. Additionally, these authors used a different population of abnormal speakers (i.e., four different communicative disorders and a control group), which also may explain the difference of the age effect in the two studies.

Past research has indicated that the sex of a listener may affect perception of a speaker's voice. Deal and Oyer (1991) found that female participants rated individuals who stuttered more positively than the male participants. Similarly, Williams and Dietrich (2001) found that their male study participants rated speakers with communication disorders as more stressed than the female participants. One reason for the sex discrepancy between the Williams and Dietrich study and the present study may be that the listeners in the current study were recruited from various backgrounds, while those in Williams and Dietrich's study were all in undergraduate programs. The diversity of backgrounds in the present study may have diluted any differences related to sex.

Acoustic data versus perceptual data

Acoustic data can be a useful tool to complement perceptual data as they enable clinicians to appreciate all aspects of a client's voice. The two surgical treatment groups in this study were rated lower in terms of the social

perception of their speech compared to both the non-surgically treated group (RT) and the control group. While the RT group was rated lower on most characteristics than the control group, these differences were not significant. The perceptual data in the present study correspond to the acoustic data from a previous study: Eksteen et al. (2003) found significant acoustic differences between the surgical and non-surgical groups. The surgical groups had a higher noise-to-harmonics ratio in addition to higher jitter and shimmer values in comparison to the control group and the RT group.

Other studies have shown that when individuals are judged as perceptually different, their acoustic data also differ from normal speakers (Zacharek et al., 2001). Dworkin et al. (2003) found that both SCL and TL subjects showed higher jitter and shimmer values in addition to lower harmonics-to-noise measures than normal. The TL and SCL groups were perceptually evaluated as having a hoarse and strained vocal quality, which is an unfortunate side effect of the surgical technique.

In contrast to individuals treated surgically, those having undergone RT for laryngeal cancer were found to display inconsistent vocal outcomes. Acoustically and perceptually, their voices are found to be either comparable to normal speakers or just slightly atypical (Dagli, Mahieu, & Festen, 1997; Aref et al., 1997). Aref et al. (1997) found that some of their RT participants exhibited abnormal values for jitter, shimmer, and harmonics-to-noise ratio, however, others displayed values within the normal range. When judged on a 7-point scale, where 1 was deemed normal and 7 was most deviant from normal, the RT subjects received a mean score of 2.4, which indicates a slightly different vocal quality. Thus, the relationship between acoustic data and perceptual data for individuals having undergone RT for laryngeal cancer is less clear.

Limitations and Future Research

The adjectives chosen for the social perception rating scale were a potential limitation in this study. As definitions were not provided to the participants, the meanings they attributed to the words may have been subjective. This lack of control may have led to difficulty in the interpretation of variables such as *boring*. Other adjectives that may have posed concerns for listeners were *clever*, *sophisticated*, and *attractive*. Qualitative feedback from the participants following the study revealed that it was difficult for some male listeners to objectively rate male speakers on the *attractive* characteristic. Additional feedback indicated that the adjectives *clever* and *sophisticated* were not easily distinguished. Furthermore, while inter-rater reliability was consistently high, intra-rater reliability was poor at times. This suggests that listeners may have wandered in their interpretation of the descriptors as they progressed through the task.

Rather than providing speakers with two normal female sample speakers, at least one male sample speaker could have been presented to help set their personal rating system. If a disordered voice had been included in addition

to male and female sample recordings, variables such as *boring* might have followed the same pattern as the other variables.

Future research in this topic area should include voice samples from female speakers to provide a more complete picture. Future studies could also include actual interactions between listeners and speakers (telephone conversations or face-to-face interactions) to make the research setting less contrived. Participants in this study frequently asked whether the patient voices they heard were computer generated and were surprised to learn that the voices were from real speakers. By including actual interactions, listeners would have an opportunity to appreciate the person behind the voice disorder.

Conclusion

Currently, there are several treatment options for laryngeal cancer. One non-surgical option is radiation therapy, while total laryngectomy and supracricoid laryngectomy involve surgical removal of substantial parts or the entire larynx. In the pre-operative counseling, information regarding the functional outcomes of each treatment should be provided to allow the patient to make an informed decision. The present study found clear differences between the surgical and the non-surgical groups. RT was rated more positively when compared to the other treatment options. The data did not indicate that individuals with TL were perceived less favorably than those with SCL. The findings may serve to provide patients and clinicians with some first orienting information about the potential social perceptual consequences of different treatment modalities for laryngeal cancer.

References

- Aref, A., Dworkin, J., Devi, S., Denton, L., & Fontanesi, J. (1997). Objective evaluation of the quality of voice following radiation therapy for T1 glottic cancer. *Radiotherapy & Oncology*, 45, 149-153.
- Blom, E. D., & Singer, M. I. (1995). Tracheoesophageal puncture and voice prosthesis for post-laryngectomy voice rehabilitation. In E. Lauder (Ed), *Self-help for the laryngectomee* (pp. 97-101). Lauder Enterprises Inc.: San Antonio, Texas.
- Blood, G. W., Mahan, B. W., & Hyman, M. (1979). Judging personality and appearance from voice disorder. *Journal of Communication Disorders*, 12, 63-67.
- Bloom, K., Zajac, D. J., & Titus, J. (1999). The influence of nasality of voice on sex-stereotyped perceptions. *Journal of Nonverbal Behavior*, 23, 271-281.
- Canadian Cancer Society/National Cancer Institute of Canada: *Canadian Cancer Statistics 2008*.
- Coman, W. B., Grigg, R. G., Tomkinson, A., & Gallagher, R. M. (1998). Supracricoid laryngectomy: a significant advance in the management of laryngeal cancer. *Australian & New Zealand Journal of Surgery*, 68, 630-634.
- Dagli, A. S., Mahieu, H. F., & Festen, J. M. (1997). Quantitative analysis of voice quality in early glottic laryngeal carcinomas treated with radiotherapy. *European Archives of Oto-Rhino-Laryngology*, 254, 78-80.
- Deal, L. V., & Oyer, H. J. (1991). Ratings of vocal pleasantness and the aging process. *Folia Phoniatrica*, 43, 44-48.
- Dworkin, J. P., Meleca, R. J., Zacharek, M. A., Stachler, R. J., Pasha, R., Abkarian, G. C., et al. (2003). Voice and deglutition functions after the supracricoid and total laryngectomy procedures for advanced stage laryngeal carcinoma. *Otolaryngology - Head & Neck Surgery*, 129, 311-320.
- Eksteen, E. C., Rieger, J., Nesbitt, M., & Seikaly, H. (2003). Comparison of voice characteristics following three different methods of treatment for laryngeal cancer. *The Journal of Otolaryngology*, 32, 250-253.
- Farrag, T. Y., Koch, W. M., Cummings, C. W., Goldenberg, D., Abou-Jaoude, P. M., Califano, J. A., et al. (2007). Supracricoid laryngectomy outcomes: The Johns Hopkins experience. *Laryngoscope*, 117(1), 129-32.
- Fietkau, R., & Sauer, R. (1992). The potentials and limits of the radiotherapy of laryngeal carcinoma. *Strahlentherapie und Onkologie*, 168, 1-16.

- Finizia, C., Dotevall, H., Lundström, E., & Lindström, J. (1999). Acoustic and perceptual evaluation of voice and speech quality: A study of patients with laryngeal cancer treated with laryngectomy vs irradiation. *Archives of Otolaryngology-Head & Neck Surgery*, 125, 157-163.
- Hurst, M. I., & Cooper, E. B. (1983). Employer attitudes toward stuttering. *Journal of Fluency Disorders*, 8, 1-12.
- Lallh, A. K., & Rochet, A. P. (2000). The effect of information on listeners' attitudes toward speakers with voice or resonance disorders. *Journal of Speech, Language, and Hearing Science*, 43, 782-795.
- Linville, S. E. (1996). The sound of senescence. *Journal of Voice*, 10, 190-200.
- McKinnon, S. L., Hess, C. W., & Landry, R. G. (1986). Reactions of college students to speech disorders. *Journal of Communication Disorders*, 19, 75-82.
- Pindzola, R. H., Auburn, A., & Cain, B. H. (1988). Acceptability ratings of tracheoesophageal speech. *Laryngoscope*, 98, 394-397.
- Qi, Y., & Weinberg, B. (1995). Characteristics of voicing source waveforms produced by esophageal and tracheoesophageal speakers. *Journal of Speech and Hearing Research*, 38, 536-548.
- Rieger, J., Dickson, N., Lemire, R., Bloom, K., Wolfaardt, J., Wolfaardt, U., et al. (2006). Social Perception of Speech in Individuals with Oropharyngeal Reconstruction. *Journal of Psychosocial Oncology*, 24 (4), 33-51.
- Schindler, A., Favero, E., Nudo, S., Spadola-Bisetti, M., Ottaviani, F., & Schindler, O. (2005). Voice after supracricoid laryngectomy: subjective, objective and self-assessment data. *Logopedics, Phoniatrics, Vocology*, 30(3-4), 114-9.
- Smith, E. S., Verdolini, K., Gray, S., Nichols, S., Lemke, J., Barkmeier, J., et al. (1996). Effect of voice disorders on quality of life. *Journal of Medical Speech-Language Pathology*, 4, 223-244.
- Starmer, H. M., Tippett, D. C., & Webster, K. T. (2008). Effects of laryngeal cancer on voice and swallowing. *Otolaryngologic Clinics of North America*, 41(4), 793-818.
- van As, C. J., Hilgers, F. J. M., Verdonck-de Leeuw, I. M., & Koopmans-van Beinum, F. J. (1998). Acoustical analysis and perceptual evaluation of tracheoesophageal prosthetic voice. *Journal of Voice*, 12, 239-248.
- van As, C. J., Koopmans-van Beinum, F. J., Pols, L. C., & Hilgers, F. J. M. (2003). Perceptual evaluation of tracheoesophageal speech by naive and experienced judges through the use of semantic differential scales. *Journal of Speech, Language, and Hearing Science*, 46, 947-959.
- Williams, D. F., & Dietrich, S. (2001). Perceptions of communicative disorders: Verification and specification of rater variables. *Journal of Communication Disorders*, 34, 355-366.
- Witt, P. D., Miller, D. C., Marsh, J. L., Muntz H. R., Grames, L. M., & Pilgram, T. K. (1997). Perception of postpalatoplasty speech differences in school-age children by parents, teachers, and professional speech pathologists. *Plastic and Reconstructive Surgery*, 100, 1655-1663.
- Zacharek, M. A., Pasha, R., Meleca, R. J., Dworkin, J. P., Stachler, R. J., Jacobs, J. R., et al. (2001). Functional outcomes after supracricoid laryngectomy. *Laryngoscope*, 111, 1558-1564.

Author Note

This research was supported by the Caritas Health Research Fund and the Alberta Heritage Foundation for Medical Research.

Correspondence concerning this article should be addressed to Jana Rieger, PhD, iRSM, Misericordia Hospital, 16940-87 Avenue, Edmonton, AB, Canada T5R 4H5. Email: jana.rieger@ualberta.ca

Received: November 5, 2008

Accepted: January 21, 2009

