

Translation, Cultural Adaptation, and Preliminary Validation of the French Version of the Thyroidectomy-Related Voice and Symptom Questionnaire (TVSQ-FR)

*Timothy Pommée, *Dominique Morsomme, *Elisa Spanoghe, †Lionel Lejeune, †‡Camille Finck, and *§Angélique Remacle, *†Liège, ‡Mons, and §Brussels, Belgium

Summary: Background. Thyroidectomy frequently leads to transient or persistent voice, swallowing, and throat or neck-related symptoms that can substantially affect patients' quality of life. Patient-reported outcome measures (PROMs) are essential to capture these subjective experiences, yet no validated French-language instrument specifically targets thyroidectomy-related voice and symptom complaints.

Objectives. This study aimed to translate, culturally adapt, and preliminarily validate a French version of the Thyroidectomy-related Voice and Symptom Questionnaire (TVSQ-FR) for use in French-speaking clinical and research settings.

Methods. Translation and cultural adaptation followed internationally recognized guidelines (WHO, ISPOR, and Cruchinho et al). Two independent forward translations were synthesized through a Delphi consensus process involving otorhinolaryngologists, speech-language pathologists, and a psychometrician, followed by independent back-translation. The prefinal version was pretested in 30 adult French-speaking patients undergoing thyroidectomy, assessed preoperatively and 2 to 4 weeks postoperatively. Quantitative analyses included internal consistency (Cronbach's α), item-total correlations, and sensitivity to postoperative change (Wilcoxon signed-rank tests). Qualitative data from open-ended questions and semistructured cognitive interviews explored comprehensibility, cultural relevance, and acceptability.

Results. Internal consistency was excellent for the Voice subscale ($\alpha = 0.90$) and total score ($\alpha = 0.92$), and good for the Throat and Neck subscale ($\alpha = 0.87$). Item-total correlations were moderate to strong across both domains. Significant postoperative changes were observed primarily in the Voice subscale, supporting the instrument's sensitivity to early postoperative symptoms. Qualitative analyses confirmed high comprehensibility, cultural adequacy, and clinical relevance, with patients describing the questionnaire as clear, accessible, and reflective of their postoperative experience.

Conclusions. The TVSQ-FR demonstrates strong linguistic and cultural equivalence with the original instrument and shows good preliminary psychometric properties. It provides a brief, patient-centered tool for assessing voice-related and throat or neck-related symptoms after thyroidectomy in French-speaking populations. Further studies with larger samples are warranted to confirm structural validity, test-retest reliability, and convergent validity, and to establish reference values for score interpretation.

Key Words: Thyroidectomy– Voice disorders– Questionnaire translation– Patient-reported outcomes– Psychometrics– French.

INTRODUCTION

Thyroidectomy is one of the most frequently performed endocrine surgeries worldwide, typically indicated for benign or malignant disorders of the thyroid gland.¹⁻⁴ In Belgium, according to the 2022 report of the National Institute for Health and Disability Insurance (INAMI), an

average of 5696 thyroidectomies are performed each year, with a median patient age of 56 years and 74.6% of cases occurring in women.⁵

Despite advances in surgical technique and intraoperative nerve monitoring, postoperative complications involving the voice, swallowing, and neck region remain common.^{2,6-8} In about 10% of cases, thyroid surgery leads to vocal fold mobility disorders, with up to 3% being permanent.^{9,10} This complication occurs because of the close anatomical relationship between the thyroid gland and the recurrent laryngeal nerve,¹¹ which innervates the intrinsic muscles of the larynx responsible for vocal fold movement. As a result, it is at risk of being stretched, compressed, or accidentally transected during surgical manipulation.¹² In addition, the external branch of the superior laryngeal nerve, which controls cricothyroid muscle tension and thus pitch modulation, can also be injured when dissecting near the superior thyroid artery, further affecting vocal function even when the recurrent laryngeal nerve is preserved.¹³

Accepted for publication February 23, 2026.

From the *Research Unit for a Life-Course Perspective on Health and Education - RUCHE, University of Liège, Liège, Belgium; †Department of Otorhinolaryngology-Head and Neck Surgery, CHU de Liège, Université de Liège, Liège, Belgium; ‡Language Sciences and Metrology Unit, Faculty of Psychology and Education, University of Mons (UMons), Mons, Belgium; and the §Center for Research in Cognition and Neurosciences, Université Libre de Bruxelles, Brussels, Belgium.

Address correspondence and reprint requests to Timothy Pommée, Research Unit for a Life-Course Perspective on Health and Education - RUCHE, University of Liège, Belgium. E-mail: timothy_pommee@hotmail.com
Journal of Voice, Vol xx, No xx, pp. xxx-xxx
0892-1997

© 2026 The Authors. Published by Elsevier Inc. on behalf of The Voice Foundation. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).<https://doi.org/10.1016/j.jvoice.2026.02.043>

Even in the absence of apparent laryngeal nerve injury, 16 to 42% of patients report voice symptoms such as voice quality changes, vocal fatigue, and reduced pitch range^{14–18}; these symptoms are most often temporary, but they can persist in about 13 to 35% of cases.¹⁵ Swallowing impairments are also very commonly reported,¹⁹ by 50–80% of patients in early weeks after surgery,²⁰ with up to 15–20% of patients showing persisting dysphagia symptoms.^{21–23} Other postoperative symptoms in the throat and neck region include sore throat, throat tightness or sensation of a throat obstruction, coughing, sensation of a foreign body in the pharynx, odynophagia, and general discomfort in the neck and perilaryngeal area.^{24–27}

The human voice plays a central role in communication, social interaction, and identity expression. Swallowing function—in addition to being a vital function for nutrition and airway protection—also contributes to comfort and social participation.^{23,28} Consequently, postoperative dysphonia and dysphagia may affect not only instrumental (eg, acoustic and aerodynamic parameters) but also psychological well-being, self-esteem, and professional functioning.^{29–31} In 2018, Kletzien et al conducted a study comparing patient-reported complaints after thyroidectomy with objective vocal parameters.³⁰ They observed a notable discrepancy between these two types of data: although most acoustic and aerodynamic measures, as well as scores on standardized voice scales, tended to return to baseline, many patients still reported persistent communication difficulties related to their voice quality. Several years postoperatively, these patients continued to experience vocal fatigue and changes in voice quality that significantly impacted their quality of life. A large body of studies has confirmed persistent throat and neck symptoms after thyroidectomy that significantly affect both personal (eg, challenges in intimate relationships) and professional life aspects (eg, occupational communication and job loss),²⁹ even when laryngeal examinations are normal.³² Beyond complaints related to phonation and swallowing, patients frequently report body image disturbances associated with weight loss, decreased sexual desire, altered sensory perception, and changes in social relationships. Many also experience pain or discomfort in the jaw, mouth, throat, and neck, as well as anxiety about their future health, all of which can affect overall quality of life.³²

These findings underscore the importance of comprehensive assessment tools that capture the patient's subjective experience rather than relying solely on objective voice measures.

Several patient-reported outcome measures (PROMs) exist for voice disorders—such as the Voice Handicap Index (VHI),³³ the Voice Symptom Scale (VoiSS),³⁴ or the Vocal Tract Discomfort Scale (VTD)³⁵—as well as for swallowing disorders—such as the Swallowing Quality of Life Questionnaire (SWAL-QOL),³⁶ the Sydney Swallow Questionnaire (SSQ),³⁷ the Eating Assessment Tool (EAT-10),³⁸ the Dysphagia Handicap Index,³⁹ or the M.D. Anderson Dysphagia Inventory (MDADI).⁴⁰ Translations to French have been done for the VHI,^{41,42} the SWAL-

QOL,^{43,44} and the SSQ⁴⁵; an original French questionnaire, the Deglutition Handicap Index, was also developed by Woisard and Lepage.⁴⁶

However, these tools were developed for general dysphonia and dysphagia populations and may not adequately reflect the specific postoperative experiences of thyroidectomy patients. To address this limitation, the first thyroidectomy-related questionnaire was developed in 2012 by Nam et al, who created a preliminary 20-item Korean tool: the Thyroidectomy-related Voice Questionnaire (TVQ).^{47,48} This questionnaire was designed to identify potential laryngeal dysfunctions before surgery and to guide further preoperative voice assessment; it includes questions about vocal complaints, dysphagia-, laryngopharyngeal reflux-, and vocal fold mobility-related symptoms associated with thyroidectomy. However, this version was never fully validated.⁴⁹ A decade later, this instrument was refined by Hwang et al (2020), resulting in the Thyroidectomy-related Voice and Symptom Questionnaire (TVSQ)—in Korean, translated to English—a psychometrically validated, 20-item self-administered tool assessing both perceived voice changes and local throat or neck discomfort.⁴⁹ The TVSQ consists of 20 items divided into two subscales: Voice (10 items) and Throat and Neck (10 items). Each item is rated on a five-point Likert scale (0 = never to 4 = always), yielding a total score range from 0 to 80, with higher scores indicating more severe symptoms. It can be used in both pre- and postoperative phases to detect preexisting or surgery-induced laryngeal dysfunctions. Its items were derived from common postoperative complaints and adapted from existing questionnaires such as the VHI, VTD, DHI, the Reflux Symptom Index (RSI),⁵⁰ and the Vocal Fatigue Index (VFI).⁵¹ The TVSQ specifically focuses on voice-related symptoms and local throat and neck discomfort, which are among the most frequent, clinically relevant, yet often underreported functional complaints following thyroidectomy.^{14,15,26} While thyroid surgery may also be associated with broader endocrine, metabolic, or systemic symptoms, these fall outside the scope of the instrument. Rather than providing a comprehensive assessment of all thyroid-related health concerns, the TVSQ is designed to capture the patient-reported functional and somatosensory consequences of thyroidectomy that directly affect voice production, swallowing comfort, and neck sensation. The validation study of the TVSQ⁴⁹ involved 50 thyroidectomy candidates, 50 patients with laryngeal disease without thyroid cancer, and 30 controls, with assessments conducted 1 week before and 2 and 4 weeks after surgery, alongside acoustic, perceptual, and aerodynamic measures. The results demonstrated high test-retest reliability, concurrent, and external validity. Kim et al⁵² confirmed its clinical sensitivity, observing that patients continued to report voice and throat/neck discomfort on the TVSQ up to 2 years post thyroidectomy, despite normal acoustic parameters. These findings suggest subtle residual laryngeal changes undetectable through laryngoscopy.

Subsequent cross-cultural adaptations have further supported the TVSQ's psychometric robustness. The Arabic,⁵³ Greek,⁵⁴ and Persian⁵⁵ versions achieved good internal consistency (Cronbach's $\alpha > 0.80$) and moderate-to-strong test-retest reliability (ICC of 0.70 for the Arabic version, 0.83 for the Persian version), confirming the tool's applicability across linguistic and cultural contexts. Nevertheless, no validated French translation is currently available, limiting its use in clinical and research settings in Belgium, France, Canada, and other Francophone regions.

Given the high demand for thyroid surgery in French-speaking populations^{5,56–59} and the lack of available tools to systematically evaluate voice and related symptoms following thyroidectomy, there is an urgent need for a culturally and linguistically validated French version of the TVSQ. Such an instrument would also facilitate international data comparison and contribute to evidence-based management of postoperative dysphonia.

The present study therefore aimed to translate, culturally adapt, and preliminarily validate the French version of the TVSQ following internationally recognized guidelines for instrument translation and adaptation.^{60–62} We hypothesized that the French TVSQ would demonstrate:

- conceptual and linguistic equivalence with the original version;
- good face validity and acceptability, reflected by high patient-reported comprehensibility and perceived relevance of the items;
- adequate internal consistency for its “Voice” and “Throat and Neck” subscales; and
- sensitivity to voice- and symptom-related changes after thyroidectomy.

METHODS

Study design and guidelines

This study followed a cross-sectional, mixed-method design including both quantitative and qualitative analyses. The translation process was informed by three authoritative guideline frameworks: WHODAS,⁶² Cruchinho et al's methodological guidelines on translation, cross-cultural adaptation, and validation of measurement instruments,⁶⁰ and the Professional Society for Health Economics and Outcomes Research (ISPOR) recommendations.⁶¹ From WHODAS, we adopted the focus on conceptual (rather than literal) equivalence, clear and patient-oriented language, attention to problematic or culturally sensitive terms, and the structured sequence of forward translation, expert review, and independent back-translation. Cruchinho et al's review provided a meta-framework emphasizing the importance of expert consensus methods (including Delphi techniques), bicultural translators, multistep adaptation procedures, and pretesting via cognitive interviews to detect cultural or methodological biases. The ISPOR guidelines further informed our use of two independent forward translations, a formal reconciliation step, an independent back-translation conducted blind to the original instrument,

and systematic cognitive debriefing with patients to assess comprehension and relevance. Integrating these elements was intended to promote strong conceptual, semantic, and cultural equivalence with the original instrument while adhering to international best practices.

Ethical considerations

The study protocol was approved by the Ethics Committee of the University Hospital of Liège (CHU de Liège; approval number: B7072024000095; reference note: 2024/257) and complied with the ethical principles of the Declaration of Helsinki. All participants provided written informed consent before inclusion. Written authorization for the translation of the TVSQ was obtained by the last author of the original publication, Prof. Dong-Il Sun, via email.

Participants

Expert committee (Delphi phase)

The Delphi consensus involved a multidisciplinary panel of experts, including otorhinolaryngologists, speech-language pathologists, and a psychometrician. Inclusion criteria were:

- native proficiency in French;
- professional experience in at least one of the following domains:—clinical or research practice related to voice or swallowing disorders—psychometrics, questionnaire development, or cross-cultural adaptation of clinical assessment tools;
- familiarity with cross-cultural adaptation of clinical tools.

Experts were excluded if they had participated in the forward translation phase to avoid bias.

Pretest participants

The pretest phase included adult patients consulting at the University Hospital of Liège for a voice assessment prior to thyroidectomy. Inclusion criteria were:

- adults (≥ 18 years old);
- native French speakers;
- scheduled for partial or total thyroidectomy.

Exclusion criteria were:

- preexisting laryngeal or neurological disorders affecting voice or swallowing;
- previous thyroid or neck surgery;
- inability to complete the questionnaire or provide informed consent.

Translation and cultural adaptation procedure

The translation procedure of the TVSQ into French can be summarized in four major steps (see [Figure 1](#)): forward translation, Delphi-based expert synthesis, back-translation, and pretesting with preliminary psychometric evaluation.

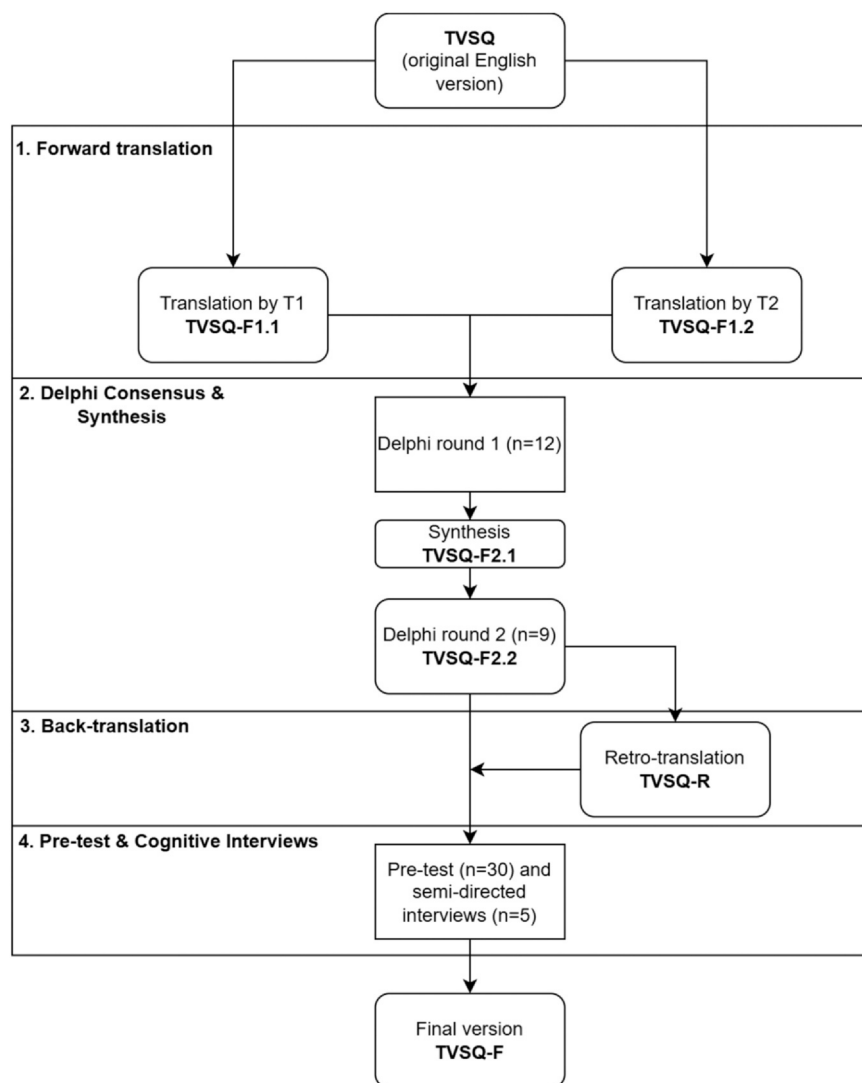


FIGURE 1. Translation steps of the TVSQ into French. TVSQ, Thyroidectomy-related Voice and Symptom Questionnaire.

Step 1: forward translation

Two independent native French-speaking researchers in the field of voice (TP and DM) produced separate forward translations of the original English TVSQ. Both were fluent in both English and French and familiar with the terminology of voice and swallowing disorders. DM is a professor and clinical speech-language pathologist (SLP) in Belgium; TP, originally trained in Belgium and France, was stationed in Canada and therefore also accustomed to the North-American cultural context.

Both translators were instructed to

- translate the entire questionnaire;
- prioritize conceptual over literal translation;
- use a natural language that is adapted to the target population;
- take into consideration potential cultural differences between anglophone and francophone patient populations;
- maintain the same level of detail for each item;

- independently translate the questionnaire without consulting each other.

Step 2: Delphi consensus survey and synthesis

To synthesize the two forward translations, a Delphi consensus procedure was conducted following established guidelines for expert consensus.^{63–67} The Delphi technique allows iterative, quasi-anonymous (participants' identities are not known to each other, but are known to the investigator) consultation of experts until a stable consensus is reached, with aggregated feedback after each round. This process minimizes peer pressure and ensures balanced participation. Items were accepted when at least 75% of experts agreed on their adequacy—a threshold defined *a priori*.

The process comprised two iterative online rounds:

- In Round 1, experts independently compared both French forward translations (TVSQ-F1.1 and TVSQ-

F1.2) with the original English items. For each item, they indicated their preferred version and provided open-ended comments regarding clarity, cultural relevance, and conceptual equivalence.

After Round 1, the research team (ES, AR, TP, and DM) reviewed quantitative preference data and qualitative feedback to draft an integrated version (TVSQ-F2.1).

In Round 2, this synthesized version was resubmitted to the panel for approval and final comments.

Free-text comments were qualitatively analyzed to identify recurring concerns (terminology, tone, and cultural nuances). The final consensus reached 88.7% agreement across all items, concluding the Delphi process and producing version TVSQ-F2.2 for back-translation.

Step 3: back-translation

The synthesized French version (TVSQ-F2.2) was back-translated into English by a licensed bilingual translator with no prior exposure to the original instrument. The back-translated version was reviewed by the research team and compared with the original to verify conceptual consistency and semantic fidelity. No modifications to the TVSQ-F2 were necessary after review of the back-translation.

Step 4: pretesting and cognitive interviews

The TVSQ-F2.2 was administered (all items were read aloud by ES) to 37 patients at two time points: preoperatively (T1) and 2 to 4 weeks postoperatively (T2). Seven patients did not participate in the post-op assessment and were thus also excluded from this study.

After the post-op administration of the TVSQ-F2.2, patients were also encouraged to respond to three open-ended questions: (1) “Do you believe you have understood all of the questions?” (2) “Would you have phrased some questions differently?” (3) “What do you think of this questionnaire?”

Five participants were further invited to take part in a semistructured cognitive interview exploring four themes: (1) comprehension, (2) opinions about the questionnaire, (3) vocal and associated symptoms, and (4) suggestions for improvement. The five patients were selected based on their responses to the TVSQ-F2.2 questionnaire, to include patients that were in the low, mid, and high score range. Interviews were conducted at home, at the patient’s workplace or online, depending on individual preference. Interviews were transcribed and verbatims were analyzed thematically.

Statistical analysis

Data were analyzed using Jamovi 2.6.22.⁶⁸ Normality was tested using the Shapiro-Wilk test. Descriptive statistics (means and standard deviations or medians and interquartile range, as appropriate) were computed for each

item and subscale of the TVSQ-FR as well as for the total TVSQ-FR score.

Internal consistency was evaluated on preoperative data using Cronbach’s alpha (α), with the following benchmarks: $\alpha \geq 0.70$ acceptable, ≥ 0.80 good, and ≥ 0.90 excellent.^{69,70} To further assess whether each item contributed appropriately to its subscale and measured the same underlying construct, we examined item-total correlations (Spearman’s rho), interpreting coefficients < 0.40 as weak, $0.40-0.69$ as moderate, $0.70-0.89$ as strong, and ≥ 0.9 as very strong.⁷¹

Sensitivity to change was tested via Wilcoxon signed-rank tests comparing pre- and postoperative scores. Rank Biserial Correlation $r^{72,73}$ was applied to demonstrate the effect size and is interpreted as follows: 0.1 indicates a small effect, 0.3 a medium effect, and 0.5 a large effect. To characterize postoperative symptom evolution, we computed individual change scores ($\Delta = \text{postoperative} - \text{preoperative}$) for the TVSQ Voice and Throat and Neck subscores. Participants were then categorized into one of three symptom-evolution profiles: improving, stable, or worsening. In the absence of established minimal clinically important difference (MCID) values for the TVSQ, improvement or worsening was defined using a distribution-based approach. An absolute change of ≥ 4 points, corresponding to the 75th percentile of the change-score distribution for each subscale in the present sample, was used as an exploratory threshold to categorize postoperative change. Profiles were then coded as (a) concurrent improvement/worsening on both subscores, (b) isolated change on a single subscore, or (c) mixed profiles showing opposite-direction changes across subscores. Given the small sample size, profiles were examined qualitatively to explore potential links between postoperative symptom evolution and etiology. Finally, correlations between symptom-change scores (total and subscores) and the number of days since surgery were computed.

Concurrent validity was preliminarily explored by analyzing GRBAS and VHI deltas, and computing correlations between TVSQ-FR (total and subscores), VHI, and GRBAS.

Qualitative data from the three open-ended questions ($n = 30$) and from the five individual interviews were analyzed thematically⁷⁴ to identify patterns related to item clarity, comprehension, and cultural relevance. A readability score was also calculated for each item using the Scolarius platform,⁷⁵ based on word, sentence, and paragraph length.

RESULTS

Delphi consensus

A panel of 12 experts (nine SLPs, two otorhinolaryngologists, and one psychometrician) participated in the first round of the Delphi process; nine experts participated in the second round (six SLPs, two otorhinolaryngologists, and one psychometrician).

In Round 1, agreement on preferred translations ranged from 50% to 100%, with a mean of 73%. Twelve of the 28

(43%) translated elements (title, instructions, scale, section headers, 20 items, and scoring labels) reached the 75% consensus threshold; the 16 remaining elements reached a moderate agreement (50-74%).

A total of 21 questionnaire elements (incl. items, instructions, and section labels) were modified following qualitative comments (incl. six items that had reached the consensus threshold) (see Appendix A for details). First, experts identified six items containing potential response-biasing formulations inherited from the original English version. Some statements appeared to presuppose the presence of postoperative symptoms (eg, the original instruction “These statements describe your voice change and throat and neck discomfort after thyroidectomy”). To prevent influencing patient responses, the research team opted to reformulate these instructions in a more neutral way (eg, “The following statements describe sensations related to your voice and your comfort in the throat and neck area.”).

Second, experts recommended tense harmonization for two items, to ensure that the French translation remained compatible with both preoperative and postoperative administrations. The English items were perceived as implicitly postoperative (eg, “I feel like my voice has become weak”), which limited their applicability at baseline. Based on expert feedback, verb tenses were systematically adjusted to the present tense (eg, “I feel like my voice is weak”), in line with the dual-timepoint purpose of the questionnaire.

Third, for 13 items, experts noted that certain terms or expressions used in the forward translations did not fully match the vocabulary or phrasings typically used by French-speaking patients. To enhance clarity and accessibility—especially given the self-administered nature of the TVSQ—the team adopted wording that reflects common patient language while preserving conceptual equivalence. Examples include selecting “voix rauque” instead of “voix cassée” (the latter being more specific to hyphophonic voices with strain), using “sens” rather than “ressens” (more related to emotions rather than physical sensations), preferring “souvent” over “fréquemment,” adding “un poids dans la poitrine” to reflect how patients commonly describe chest pressure, and harmonizing wording and structure across items so that one item beginning with “My voice” was restructured to fit the same “I feel/I have...” format used throughout the questionnaire.

This process led to the creation of the provisional synthesized version TVSQ-F2.1 that was then submitted to the expert panel for Round 2 validation. Nine experts participated in this second round. The TVSQ-F2.1 achieved a 88.7% consensus.

After Round 2, ES, AR, TP, and DM met to review the experts' comments on the TVSQ-F2.1 and discuss any potential remaining issues. Several participants had raised ambiguity on the items 5: “J'ai des difficultés à produire une voix forte” (“I have difficulties producing a loud voice”) and 6 “J'ai l'impression que ma voix est faible” (“I

feel like my voice is weak”), relating to the concept of voice loudness. Therefore, an example was added to item 5: “J'ai des difficultés à produire une voix forte (par exemple: appeler quelqu'un dans la maison)” (“I have difficulties producing a loud voice (eg, calling someone in the house”). This meeting resulted in the final version of the TVSQ-F (see Appendix B).

Back-translation

Given the extensive discussions and revisions that had shaped our translation, it was expected that the back-translation of certain items would deviate slightly from the original wording. Nevertheless, none of these differences raised concerns or required further modification of the TVSQ-F2.

Preliminary test

Pretest sample characteristics

Thirty participants (24 women, six men; mean age = 55.2 years \pm 13.5, range 30-79) completed both pre- and post-operative assessments. The median number of days between pre-op assessment and surgery was 18 (IQR = 20, range 2-41), the median number of days between surgery and post-op assessment was 13 (IQR = 13.5, range 5-55). Median days between pre- and post-op assessment was 35 (IQR = 25.3, range 14-64). All surgeries were carried out at the University Hospital of Liège. The majority of patients underwent total thyroidectomy (n = 24/30, 80%), of which three underwent an additional parathyroidectomy. One patient underwent parathyroidectomy only, and one had a partial thyroidectomy (lobo-isthmectomy); surgical details were unavailable for four patient files.

Thyroid pathology accounted for most cases (28; 93%), predominantly dystrophic or nodular thyroid disease (18; 60%) and multinodular goiter (nine; 30%), including one plunging goiter requiring sternotomy. Hyperthyroidism to Graves-Basedow' disease or nodular autonomy accounted for four cases (13%). A parathyroid disorder was present in six patients (20%), mainly primary hyperparathyroidism, sometimes associated with concurrent thyroid disease. Common comorbidities included type II diabetes and bone demineralization related to hyperparathyroidism.

Three cases of damage to the superior laryngeal nerve during the thyroid surgery were documented.

Preoperatively, most patients presented with normal or near-normal voice quality, with the majority receiving a score of 0 on Grade (70%), Roughness (63%), Breathiness (81%), Asthenia (93%), and Strain (67%) (Table 3). Only a small minority showed mild deviations (scores 1-2), and no severe dysphonia (score 3) was observed. Preoperative VHI scores were low across all domains, with a median total score of 2 (IQR = 14), indicating minimal perceived vocal handicap at baseline. Subscale scores were similarly low (Functional: median 0; Emotional: median 0; Physical: median 2; Table 4).

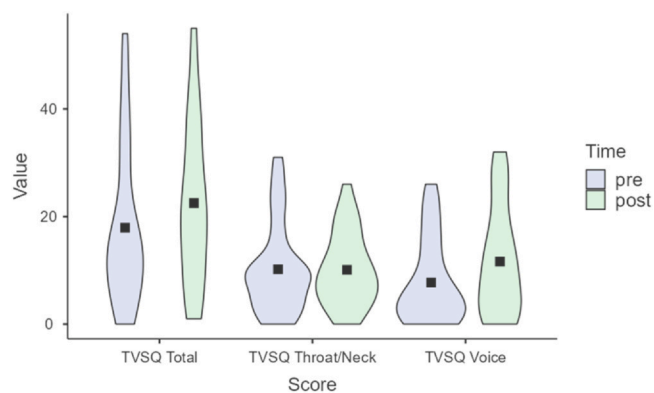


FIGURE 2. Pre- and postoperative violin plots for total and subscores of the TVSQ-FR; black square: median score; $n = 30$. TVSQ, Thyroidectomy-related Voice and Symptom Questionnaire.

TVSQ-FR: descriptive statistics. Figure 2 shows the pre- and postoperative distribution of TVSQ-FR total scores, as well as the two subscales. Figure 3 shows the score distribution pre- and post-surgery for individual items of the TVSQ-FR, grouped by subscale.

Internal consistency and item-total correlations. Internal consistency on preoperative data was excellent for the Voice subscale (Cronbach's $\alpha = 0.90$) and for the total TVSQ-FR score (Cronbach's $\alpha = 0.92$); internal consistency was good for the Throat and Neck subscale (Cronbach's $\alpha = 0.87$).

All preoperative item-total correlations were statistically significant; six items from the Voice subscale demonstrated a strong correlation (> 0.7), four showed moderate correlations; two Throat and Neck items showed a strong correlation with the subscale, the other eight showed a moderate correlation (Table 1).

Sensitivity to change. The total TVSQ-FR score and subscores showed a tendency to increase after surgery (Voice subscore: on average $+ 3.9$ points (± 8.8), $W = 83.5$, $P = 0.03$; Throat-Neck subscore: $+ 0.4$ (± 5.8), $W = 252$, $P = 0.70$; Total score: mean $+ 5.5$ (± 11.8), $W = 133$, $P = 0.07$).

Four individual items (three from the Voice and one from the Throat-Neck subscale) as well as the Voice subscore showed a statistically significant increase ($P < 0.05$, Table 2), all with a large effect size, except for the latter, which showed a medium effect size.

Postoperative symptom trajectories showed substantial interindividual variability. Some participants demonstrated worsening Voice and/or Throat and Neck symptoms, others remained stable, and a smaller number showed improvement, particularly in Throat and Neck symptoms. No consistent relationship was observed between etiology and symptom evolution. No significant correlations were observed between the magnitude of TVSQ total or subscale change scores and the time elapsed between surgery and postoperative assessment ($P > 0.05$).

Concurrent validity. Postoperatively, a shift toward greater perceived dysphonia was observed on the GRBAS scale (Table 3). The proportion of score 0 decreased mostly for G, R, and S, with increases in mild (score = 1) or moderate (score = 2) deviations, particularly for Grade and Roughness. Severe deviation (score = 3) was only observed postoperatively for Roughness ($n = 1$).

VHI total scores, which were low preoperatively, increased substantially after surgery (Table 4). This change was observed across all subscales, with an increase twice as high for the Physical subscore.

Preoperatively, no statistically significant correlation was found between TVSQ-FR scores and the GRBAS parameters ($n = 27$). The TVSQ Voice subscore was significantly correlated with the VHI total ($r_s = 0.73$, $P = 0.03$) and Physical subscore ($r_s = 0.70$, $P = 0.03$) ($n = 9$).

Postoperative concurrent validity analyses were not interpreted due to very small sample sizes ($n = 5$).

Readability and comprehension. Scolarius readability scores ranged from 33 to 81 (mean = 53, standard deviation = 12), all items corresponding to a comprehension level equivalent to primary school.

All participants ($n = 30$) reported full comprehension of the questionnaire items. Cognitive interviews ($n = 5$) further confirmed that the TVSQ-FR was clear, easy to understand, and relevant to participants' postoperative experience.

A small number of participants identified minor ambiguities in specific items, particularly regarding the distinction between vocal weakness and difficulty producing a loud voice (4/30), and the interpretation of the singing voice item (7/30). These observations reflected differences in interpretation rather than lack of comprehension and informed minor wording clarifications. A few participants also activated certain items for reasons unrelated to thyroidectomy (eg, medication-related dry mouth or allergy-related respiratory difficulties).

Comprehensiveness and relevance. Most interviewees (4/5) judged the questionnaire to be comprehensive and adapted to their situation, with items that were generally straightforward and relevant to their postoperative experience. Reported postoperative symptoms corresponded closely to existing TVSQ-FR items, supporting the instrument's content validity. However, two participants mentioned the absence of an item related to scar-related sensations. Overall, the interviewed participants appreciated the TVSQ-FR's focus on symptoms beyond voice, particularly physical sensations such as throat pressure and neck stiffness. One expressed that these issues are rarely addressed in postoperative care despite their impact on daily communication and comfort.

In addition to the specific item ambiguities already mentioned (singing voice, vocal weakness/loudness, mouth dryness, and missing scar-related item), two participants suggested including a short introductory explanation

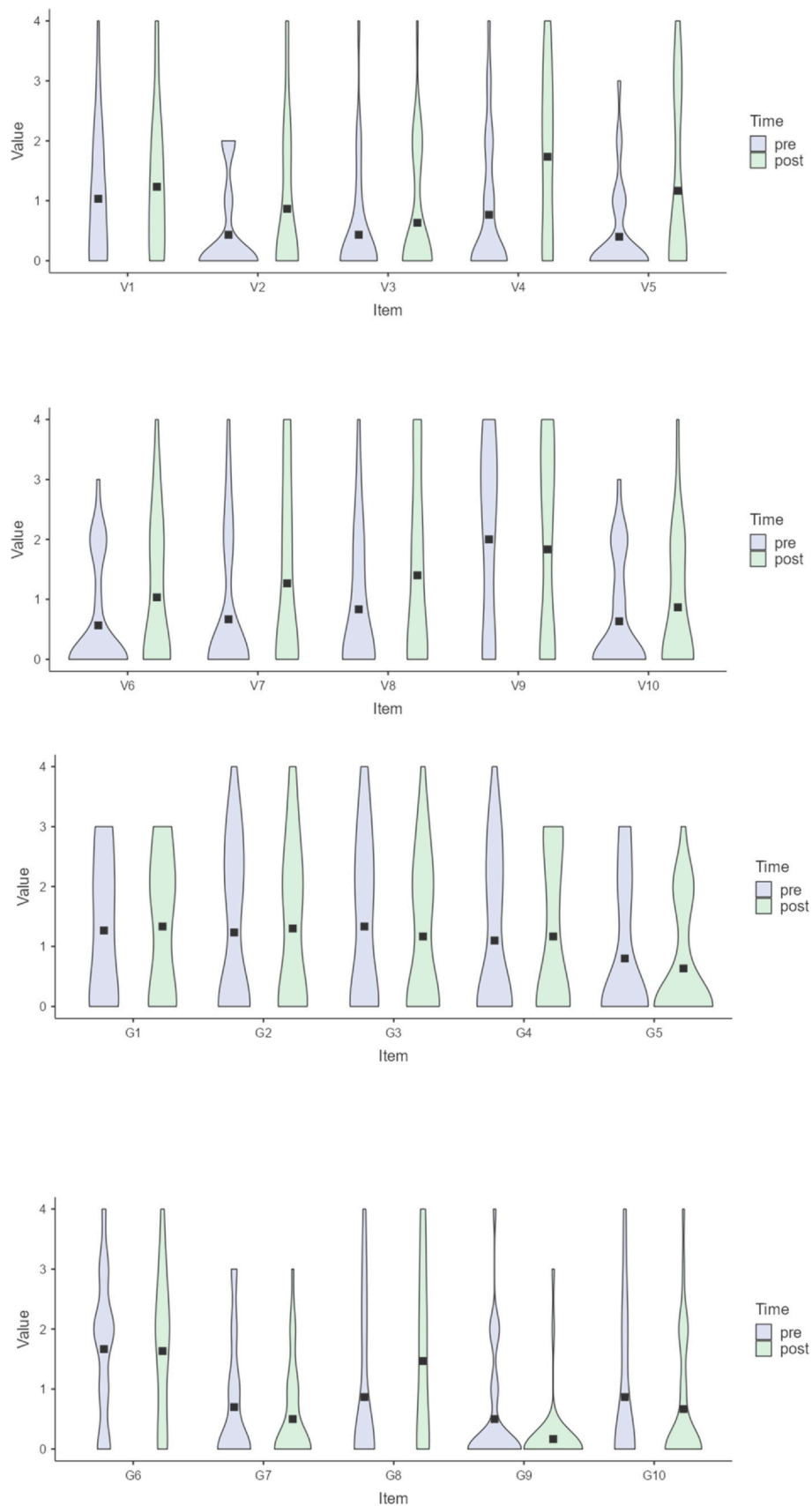


FIGURE 3. Pre- and postoperative violin plots for individual items of the TVSQ-FR (voice: top; throat-neck: bottom); black square: median score; n = 30. TVSQ, Thyroidectomy-related Voice and Symptom Questionnaire.

TABLE 1.
Item-Subtotal Correlations for the 20 TVSQ Items (Preoperatively)

Item	Voice Subscore	Item	Throat-Neck Subscore
V1	$r_s = 0.72$ ($P < 0.001$)	G1	$r_s = 0.55$ ($P = 0.002$)
V2	$r_s = 0.74$ ($P < 0.001$)	G2	$r_s = 0.87$ ($P < 0.001$)
V3	$r_s = 0.55$ ($P = 0.002$)	G3	$r_s = 0.81$ ($P < 0.001$)
V4	$r_s = 0.75$ ($P < 0.001$)	G4	$r_s = 0.58$ ($P < 0.001$)
V5	$r_s = 0.61$ ($P < 0.001$)	G5	$r_s = 0.58$ ($P < 0.001$)
V6	$r_s = 0.66$ ($P < 0.001$)	G6	$r_s = 0.51$ ($P = 0.005$)
V7	$r_s = 0.77$ ($P < 0.001$)	G7	$r_s = 0.67$ ($P < 0.001$)
V8	$r_s = 0.63$ ($P < 0.001$)	G8	$r_s = 0.65$ ($P < 0.001$)
V9	$r_s = 0.77$ ($P < 0.001$)	G9	$r_s = 0.42$ ($P = 0.022$)
V10	$r_s = 0.79$ ($P < 0.001$)	G10	$r_s = 0.54$ ($P = 0.002$)

outlining the purpose of the questionnaire and clarifying that not all listed symptoms are expected to occur.

Emotional and experiential responses. Participants' emotional responses to the questionnaire varied. Two described the experience as eye-opening, noting that completing the questionnaire made them aware of symptoms they had not fully recognized and promoted self-reflection, helping them identify changes in their voice or swallowing and encouraging greater attention to their bodily sensations. For two participants, the list of possible symptoms created initial anxiety or surprise, as they had not expected the range of postoperative experiences presented. Two found the process reassuring, noting the questionnaire helped them structure their symptoms and prepare questions for the otorhinolaryngologist to understand their recovery better.

DISCUSSION

This study presents the first French translation and cultural adaptation of the Thyroidectomy-Related Voice and Symptom Questionnaire (TVSQ-FR) and provides preliminary evidence of its psychometric validity in a French-speaking clinical population. Using a rigorous, guideline-based multistep methodology, we demonstrated that the TVSQ-FR preserves the conceptual structure of the

TABLE 3.
Pre- and Postoperative GRBAS Score Distributions

Scale	% Score 0 (Pre/Post)	% Score 1 (Pre/Post)	% Score 2 (Pre/Post)	% Score 3 (Pre/Post)
G	70%/20%	26%/60%	4%/20%	0%/0%
R	63%/20%	33%/60%	4%/0%	0%/20%
B	81%/80%	19%/20%	0%/0%	0%/0%
A	93%/80%	7%/0%	0%/20%	0%/0%
S	67%/40%	30%/40%	4%/20%	0%/0%

Notes: Preoperative n = 27; postoperative n = 5.

TABLE 4.
Pre- and Postoperative VHI Scores (Median, IQR, and Range)

	Preoperative (n = 9)	Postoperative (n = 4)
VHI total	Median = 2 IQR = 14 Range = 0-28	Median = 43 IQR = 29.5 Range = 3-61
Functional	Median = 0 IQR = 6 Range = 0-9	Median = 10.5 IQR = 15.5 Range = 0-23
Emotional	Median = 0 IQR = 2 Range = 0-9	Median = 9 IQR = 9 Range = 0-18
Physical	Median = 2 IQR = 4 Range = 0-15	Median = 21.5 IQR = 7.5 Range = 3-24

original instrument while remaining linguistically accessible, culturally appropriate, and clinically relevant. Overall, the results support the TVSQ-FR as a comprehensive and comprehensible patient-centered PROM for assessing voice-related and throat/neck symptoms in French-speaking patients undergoing thyroidectomy.

Cross-cultural adaptation and face validity

During the translation and adaptation process, several changes were introduced beyond literal translation to strengthen face validity—a crucial component of PROM acceptability. The expert review identified response-biasing formulations inherited from the English version—particularly phrasings that implicitly presupposed the presence of

TABLE 2.
Statistically Significant Pre- vs Postoperative Score Differences According to Wilcoxon (W) Test With Rank Biserial Correlation Effect Sizes (r), n = 30

Variable	W	P Value	Effect Size (r)	Pre → Post Means (SD)	Pre → Post Medians (IQR)
V4 (vocal fatigue)	16.5	0.003	-0.81	0.76 (1.14) → 1.73 (1.44)	0 (1.00) → 2 (3.00)
V5 (difficulties producing loud voice)	18	0.017	-0.70	0.40 (0.77) → 1.17 (1.46)	0 (0.75) → 0 (3.00)
V8 (difficulty producing high pitch)	14	0.028	-0.69	0.83 (1.15) → 1.40 (1.52)	0 (1.75) → 1 (2.00)
G8 (neck feels numb or uncomfortable)	19	0.036	-0.64	0.87 (1.36) → 1.47 (1.48)	0 (2.00) → 1 (2.75)
Voice subscore	83.5	0.034	-0.49	7.73 (7.83) → 11.63 (10.0)	4.5 (11.8) → 10 (14.8)

postoperative symptoms—which were reformulated into neutral descriptions. Verb tense adjustments were also required to ensure that items remained applicable in both preoperative and postoperative contexts (eg, “My voice is weak” instead of “My voice has become weak”). In addition, vocabulary harmonization allowed the wordings to reflect French-speaking patients’ natural phrasing (eg, “voix rauque” instead of “voix cassée” “sens” instead of “ressens”), while syntactic standardization reduced structural inconsistencies across items.

Similar to the Persian⁵⁵ and Arabic⁵³ adaptations of the TVSQ, cognitive interviews were used to assess comprehension and cultural relevance. Unlike the Persian study, which did not identify substantial post-debriefing revisions, our qualitative analyses revealed ambiguity between the concepts of vocal weakness (“faiblesse vocale”) and difficulty producing a loud voice (“difficulté à parler fort”). Adding a concrete example to the latter item improved interpretability without altering its conceptual intent. This illustrates the value of integrating patient feedback even when expert consensus is high.

Readability scores and qualitative feedback confirmed that the final TVSQ-FR items were easily understood and perceived as relevant. Patients consistently described the TVSQ-FR as easy to understand, accessible, comprehensive, and reflective of their postoperative experience. Spontaneous vocal and somatosensory symptom reports closely echoed the dual-domain structure of the TVSQ, reinforcing both its face validity and cross-cultural consistency.

Several participants described the questionnaire as eye-opening, helping them identify symptoms they had not previously articulated. Others found the list of symptoms initially anxiety-provoking, confronting them with the breadth of potential postoperative symptoms. These contrasting reactions illustrate how PROMs can simultaneously support self-reflection, symptom awareness, and patient engagement,^{76,77} but also trigger hypervigilance—an important factor when interpreting postoperative self-reports. Our findings underline the importance of contextualizing PROMs during clinical use. As suggested during cognitive debriefing, introducing a brief explanatory statement clarifying the purpose of the questionnaire, emphasizing that not all listed symptoms are expected, and explicitly instructing patients to respond in relation to thyroidectomy-related effects may help reduce anxiety and prevent overinterpretation of the listed symptoms. A free-text comment field could further allow patients to contextualize symptoms unrelated to surgery, such as medication-induced dry mouth or respiratory comorbidities.

Psychometric performance

Internal consistency and measurement structure

The TVSQ-FR demonstrated excellent internal consistency for both subscales, consistent with the previously published cross-cultural adaptations. Cronbach’s alpha values for the Voice and Throat and Neck domains ($\alpha = 0.90$ and 0.87 , respectively) were comparable to those reported in the Greek ($\alpha = 0.95$ and 0.85),⁵⁴ Persian ($\alpha = 0.95$ and 0.88),⁵⁵

and Arabic versions ($\alpha = 0.94$ and 0.89).⁵³ These values confirm that the questionnaire reliably captures coherent constructs and that the underlying structure of the instrument translates well across languages and clinical contexts. The item-total correlations further support internal cohesion of the scale, although—as observed in the Greek validation—throat and neck items generally show lower correlations than voice-related items.⁵⁴ This pattern likely reflects the lower specificity of throat and neck symptoms, such as neck stiffness, throat tightness, or globus sensation, which may arise from multiple sources, including surgical manipulation, endotracheal intubation, musculoskeletal tension, or stress, rather than being uniquely linked to laryngeal function. Given the modest sample size and the well-established two-factor structure of the original TVSQ, factor analysis was not conducted at this stage. Larger samples will be required to confirm structural validity in French-speaking populations.

Sensitivity to postoperative change

Prior research has shown that up to 25-90% of patients report postoperative dysphonia in the first weeks, with persistent symptoms in 15% of cases,⁷⁸ up to 2 years post surgery.⁵² Beyond dysphonia, nonvoice symptoms such as neck tightness, globus, dysphagia, and sensory changes are also common following thyroidectomy, with reported prevalence ranging from 20 to 70%^{26,79,80} depending on symptom and timing of assessment. Our findings showed significant postoperative increases mostly in the Voice subscore (especially for vocal fatigue, difficulties producing loud voice, and high pitch)—only one item from the Throat and Neck subscore (neck numbness) reached significance. This asymmetry is consistent with symptom trajectories reported in the original TVSQ validation⁴⁹ and subsequent cross-cultural Arabic adaptation.⁵³ The Persian version on the other hand showed similarly marked increases in both scores 2-week postoperative scores,⁵⁵ and the Greek study did not report pre- and postoperative scores. Hwang et al mentioned in their original validation study that “the total score of the two subcategories may increase or decrease based on a particular patient’s symptoms.” Indeed, the patterns observed in our study, combined with previous findings, highlight the importance of a separated score for these two subdomains, reflecting the multifactorial nature of postoperative morbidity, where even without nerve injury, voice production can be affected by infrahyoid muscle division, local edema, altered laryngotracheal dynamics, increased pain and cervical muscle tension and stiffness, and changes in breath support and somatosensory feedback.^{22,81,82}

Our findings also highlight substantial interindividual variability in postoperative symptom trajectories, consistent with previous literature. While prior studies have identified associations between clinicopathologic factors and postoperative symptom severity (eg, lymph node involvement⁸³ and extents of thyroidectomy and neck dissection^{52,84}), the variability observed in our cohort underscores the multifactorial nature of postoperative recovery, most likely influenced by

differences in biomechanical factors (eg, inflammatory response, postoperative muscle tension, and scar formation), compensatory vocal strategies, proprio- and interoceptive abilities, and psychosocial factors. Notably, Kim *et al*⁵² identified voice abuse history, professional voice use, and higher preoperative vocal pitch as significant predictors of elevated TVSQ score at 2 years after surgery. These findings reinforce the value of patient-reported outcome measures such as the TVSQ-FR, which capture subjective morbidity not fully explained by surgical or anatomical factors alone.

Concurrent validity

Although the limited sample size of postoperative GRBAS and VHI data does not allow firm conclusions to be drawn with regard to the TVSQ-FR's concurrent validity, the observed patterns of association across perceptual and self-reported measures suggest underlying differences in the constructs they capture.^{52,83,84} The TVSQ-FR is designed to assess situational throat and voice sensations experienced in daily life, whereas the VHI reflects a broader construct of voice-related participation restriction and perceived handicap, and the GRBAS reflects clinician-rated auditory-perceptual deviation. This distinction may help explain why, preoperatively, the TVSQ-FR Voice subscore was correlated with VHI scores but not with clinician-rated GRBAS parameters: self-reported symptoms and perceived handicap may align more closely in stable, low-severity contexts, whereas auditory-perceptual deviations remain subtle or absent. Our findings are consistent with previous validation studies showing stronger associations between the TVSQ and symptom-focused instruments such as the Vocal Tract Discomfort (VTD)³⁵ scale than with global handicap measures. The Persian validation study⁵⁵ found moderate correlations with the VHI and strong correlations with the VTD, highlighting the TVSQ's alignment with both functional and somatosensory domains of voice-related symptoms. Notably, in both the Persian and original validations, Throat and Neck items tended to correlate more strongly with VTD than with VHI, likely because VHI emphasizes communicative impact, while VTD captures physical sensations. The present French findings therefore support the construct specificity of the TVSQ-FR rather than a lack of convergent validity. Future studies with larger samples should further examine convergent and discriminant validity against multiple PROMs, such as the Voice Symptom Scale (VoiSS),³⁴ VTD, EAT-10,³⁸ and SWAL-QOL.³⁶

Clinical significance

Voice and throat and neck-related complications after thyroidectomy are often underestimated, particularly when no visible laryngeal nerve injury is present.^{14,15} Standard postoperative evaluations often prioritize acoustic measures, clinician-rated perception, and laryngoscopic findings, which may fail to capture subtle yet impactful symptoms affecting communication, comfort, and quality

of life.^{29,30} The TVSQ-FR provides a brief, accessible, and patient-centered means of addressing this gap.

Its dual focus on phonatory function and throat/neck sensations aligns with contemporary understanding of post-thyroidectomy morbidity as a multidimensional phenomenon.^{14,15} Routine use of the TVSQ-FR may facilitate early identification of patients requiring targeted voice therapy, physiotherapeutic intervention, or multidisciplinary follow-up, and may support more individualized postoperative care pathways.

Limitations and future directions

This study followed rigorous international guidelines for questionnaire translation and incorporated both expert consensus (Delphi procedure) and patient feedback. This multi-method approach ensured linguistic precision and cultural adaptation. However, the preliminary validation included in this study showed several limitations. First, the sample size was modest, limiting statistical power and precluding factor analysis as well as the generalizability of our findings; future research should aim to replicate these findings in larger French-speaking cohorts, ideally across diverse French-speaking populations—including France, Belgium, Switzerland, and Canada—to ensure cross-cultural equivalence and clinical utility. Second, the short postoperative follow-up period primarily reflects early inflammatory changes rather than stable recovery outcomes. Longer-term follow-ups will allow for assessment of the TVSQ's relevance for more stable phases of the postoperative recovery, including longitudinal studies for the evaluation of the questionnaire's responsiveness to rehabilitation. Third, the heterogeneity of surgical procedures (total vs partial thyroidectomy, with or without parathyroidectomy) reflects real-world clinical practice and supports ecological validity and broad applicability of the TVSQ across thyroid surgery patients. However, this heterogeneity may have limited the ability to examine procedure-specific symptom profiles in the present sample. Future research should evaluate discriminant validity with more homogeneous laryngeal pathology groups. Last but not least, concurrent validity analyses were constrained by incomplete postoperative GRBAS/VHI data. Further investigations should also explore test-retest reliability, objective-subjective correlations (eg, acoustic, aerodynamic, and EMG), and the establishment of reference values in French-speaking populations to support score interpretation and clinical decision-making.

CONCLUSION

This study presents the first French translation and cultural adaptation of the Thyroidectomy-related Voice and Symptom Questionnaire (TVSQ-FR). The TVSQ-FR demonstrated excellent internal consistency, good sensitivity to postoperative changes, and high patient acceptability, indicating strong preliminary psychometric validity. The combination of quantitative and qualitative evidence supports its integration into routine perioperative monitoring, interdisciplinary care, and future cross-linguistic research

on thyroidectomy outcomes. By integrating both voice-related and throat/neck-related symptoms, the TVSQ-FR provides a nuanced, patient-centered assessment of the postoperative experience—addressing dimensions often overlooked in conventional evaluations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank all patients who generously participated in this study. We also thank the members of the expert committee who contributed to the Delphi consensus process for their valuable time, expertise, and constructive feedback during the translation and cultural adaptation of the questionnaire. We are grateful to Prof. Dong-II Sun for granting permission to translate and adapt the original Thyroidectomy-related Voice and Symptom Questionnaire (TVSQ) into French. We also acknowledge the clinical teams of the University Hospital of Liège for their support with patient recruitment and data collection.

Appendix A. Documentation of the Translation and Adaptation Process

1. Front Matter (Title, Instructions, Scale, and Section Titles)

Element	Original English Wording	Favored Translation in Delphi Round 1	% of Agreement	Final Item	Rationale
Title	Thyroidectomy-Related Voice and Symptom Questionnaire	Questionnaire sur les changements vocaux et les symptômes liés à la thyroïdectomie	66.7%	Questionnaire relatif aux symptômes et à la qualité vocale dans le cadre d'une chirurgie thyroïdienne	The term "change" was considered potentially biasing, as it may prompt respondents to compare their current voice to a previous state rather than evaluate their present experience. In addition, "change" is only appropriate for post-operative administration and not for preoperative use of the questionnaire.
Instructions	These statements describe your voice change and throat and neck discomfort after thyroidectomy. Circle the response that indicates how frequently within the past month you have encountered the same experience.	Les affirmations suivantes décrivent les changements ressentis en termes de qualité vocale et/ou de gêne ressentie au niveau de la gorge et du cou après votre thyroïdectomie. Entourez la réponse qui indique la fréquence à laquelle vous avez rencontré ces changements au cours du dernier mois / Les énoncés suivants décrivent les changements vocaux ainsi que les gênes ressenties au niveau de la gorge et du cou après une thyroïdectomie. Entourez la réponse qui indique à quelle fréquence, au cours du dernier mois, vous avez rencontré l'expérience décrite.	50%	Les affirmations suivantes décrivent des ressentis en lien avec votre voix et votre confort au niveau de la gorge et du cou. Entourez le chiffre correspondant le mieux à votre expérience au cours du dernier mois.	Same rationale as for the title regarding the term "change." The term "experience" was therefore adopted as a more neutral alternative.

Element	Original English Wording	Favored Translation in Delphi Round 1	% of Agreement	Final Item	Rationale
Scoring scale	0 (never), 1 (almost never), 2 (sometimes), 3 (almost always), or 4 (always)	0 (jamais), 1 (presque jamais), 2 (parfois), 3 (presque toujours), 4 (toujours)	100%	0 (jamais), 1 (presque jamais), 2 (parfois), 3 (presque toujours), 4 (toujours)	This option received unanimous agreement among experts.
Section title 1	Voice change	Changements vocaux	66.7%	Concernant votre voix	The term “voice change” was considered potentially biasing, as it may encourage respondents to focus on perceived changes rather than current vocal status. Moreover, it is not suitable for preoperative administration. A neutral phrasing was therefore preferred.
Subscore 1	Voice change score	Score relatif à la modification de la qualité vocale	58.3%	Score relatif à la voix	Terms implying change or modification were avoided, as they may bias respondents toward comparative judgments and are inappropriate for preoperative administration.
Section title 2	Throat and neck discomfort	Gêne dans la gorge et le cou	66.7%	Concernant votre gorge et votre cou	The term “discomfort” was considered potentially suggestive. The neutral phrasing “regarding” was preferred to avoid prompting respondents to actively search for symptoms.
Subscore 2	Throat and neck discomfort score	Score relatif à la gêne au niveau de la gorge et du cou	91.7%	Score relatif à la gorge et au cou	The term “discomfort” was avoided to reduce suggestive bias. The neutral phrasing remains consistent with the first section of the questionnaire.
Total score	Total score	Score total	91.7%	Score total	This wording obtained a clear majority of expert votes.

2. Voice Items (V1-V10)

Item	Original English Wording	Favored Translation in Delphi Round 1	% of Agreement	Final Item	Rationale
V1	My voice sounds hoarse and cracked	Ma voix semble enrouée ou cassée	58.3%	Ma voix semble rauque ou enrouée	According to clinician experts, "rough" (rauque) is more commonly used by patients, whereas "cracked" (cassée) is typically associated with hypophonia following vocal overuse.
V2	I feel strained when producing my voice	J'ai l'impression que je dois forcer pour produire ma voix	91.7%	J'ai l'impression que je dois forcer pour produire ma voix	This wording obtained a clear majority consensus.
V3	I feel pain or discomfort after talking	Je ressens de la douleur ou une gêne après avoir parlé	100%	Je sens une douleur ou une gêne après avoir parlé	The verb "feel" (sentir) was preferred over "experience" (ressentir), as it is more closely associated with physical sensations rather than emotional states. "De la" was replaced by "une," which places less emphasis on the quantity of discomfort experienced.
V4	I feel vocal fatigue after a long conversation	Je ressens une fatigue vocale après une longue conversation	58.3%	Je sens une fatigue vocale après une longue conversation	Same rationale as for item V3: preference for wording associated with physical sensation.
V5	I have difficulty producing a loud voice	J'ai des difficultés à produire une voix forte	75%	J'ai des difficultés à produire une voix forte (par exemple: appeler quelqu'un dans la maison)	This version obtained the majority of expert votes. (The concrete example was later added following cognitive interviews.)
V6	I feel like my voice has become weak	J'ai l'impression que ma voix est devenue faible / J'ai l'impression que ma voix s'est affaiblie	50%	J'ai l'impression que ma voix est faible	The present tense was chosen to allow use of the item in both preoperative and postoperative contexts, avoiding an explicit comparison over time.
V7	I have difficulty singing	J'ai des difficultés pour chanter	91.7%	J'ai des difficultés pour chanter	This wording obtained a strong majority consensus.
V8	I have difficulty producing a high pitch	J'ai des difficultés à produire un son aigu	58.3%	J'ai des difficultés à produire un son aigu	This version obtained the majority of expert votes.
V9	I feel like my voice tone is lower than before	J'ai l'impression que ma voix est plus grave qu'avant / J'ai l'impression que le ton de ma voix est plus grave qu'avant.	50%	J'ai l'impression que le son de ma voix est grave	The present tense was adopted to ensure applicability in both preoperative and postoperative assessments.
V10	I run out of air when I talk	Je suis à bout de souffle quand je parle	100%	Je suis à bout de souffle quand je parle	This translation obtained unanimous expert agreement.

3. Throat and Neck Items (G1-G10)

Item	Original English Wording	Favored Translation in Delphi Round 1	% of Agreement	Final Item	Rationale
G1	I have large amounts of sputum in my throat	J'ai beaucoup de salive/mucus/glaïres dans ma gorge	66.7%	J'ai beaucoup de salive/mucus/glaïres dans la gorge	Patients more commonly use the definite article ("la") rather than the possessive form ("ma").
G2	I feel like something is stuck in my throat	J'ai l'impression d'avoir quelque chose de coincé dans ma gorge	91.7%	J'ai l'impression d'avoir quelque chose de coincé dans la gorge	Same rationale as Item G1.
G3	I frequently clear my throat because I feel that I have sputum in my throat	Je me racle fréquemment la gorge à cause de la présence de salive/mucus/glaïres	66.7%	Je me racle souvent la gorge car je sens beaucoup de salive/mucus/glaïres	The adverb "souvent" ("often") was preferred over "fréquemment" ("frequently") to better reflect patient language.
G4	I feel discomfort when swallowing food or liquid	Je ressens de la gêne lorsque j'avale des aliments ou des liquides	91.7%	Je sens une gêne lorsque j'avale des aliments ou des liquides	The verb "sentir" ("feel") was preferred for physical sensation, and the article was adjusted to avoid implying quantity of discomfort.
G5	I cough after meals or after lying down	Je tousse après le repas ou lorsque je suis allongé	66.7%	Je tousse après avoir mangé ou lorsque je suis allongé-e	This wording was considered more representative of the clinical reality described by the item. Inclusive writing added.
G6	My mouth or throat is dry and I feel thirsty	Ma bouche ou ma gorge est sèche et j'ai soif	58.3%	J'ai la bouche ou la gorge sèche, et j'ai soif	Use of first-person singular harmonizes phrasing across items.
G7	I have difficulty breathing or frequent choking episodes	J'ai des difficultés à respirer ou j'ai fréquemment des épisodes d'étouffement / J'éprouve des difficultés pour respirer ou j'ai des épisodes d'étouffement fréquents	50%	J'ai des difficultés respiratoires ou j'ai fréquemment des épisodes d'étouffement	The phrase "difficultés respiratoires" was judged to better reflect clinical usage.
G8	My neck is numb and I feel discomfort (or pain)	Mon cou est engourdi et je ressens une gêne (ou de la douleur)	91.7%	Mon cou est engourdi et je sens une gêne (ou une douleur)	Preference for wording associated with physical sensation ("sentir").
G9	My upper chest is numb and I feel discomfort (or pain)	Je sens un poids sur ma poitrine, une gêne voire une douleur / Ma poitrine est engourdie et je ressens une gêne (ou de la douleur).	50%	Je sens un poids, une gêne ou une douleur dans la poitrine	Patients more commonly report a sensation of "pressure in the chest" that also better reflects the intended clinical construct.
G10	My shoulder is numb and I feel discomfort (or pain)	Mes épaules sont engourdies et je ressens une gêne (ou de la douleur)	91.7%	Mes épaules sont engourdies et je sens une gêne (ou une douleur)	Same rationale as Items G3 and G8 and the article "de la" was adjusted to avoid implying quantity of discomfort.

Appendix B. Final Version of the TVSQ-FR

Questionnaire relatif aux symptômes et à la qualité vocale dans le cadre d'une chirurgie thyroïdienne

Les affirmations suivantes décrivent des ressentis en lien avec votre voix et votre confort au niveau de la gorge et du cou. Entourez le chiffre correspondant le mieux à votre expérience au cours du dernier mois.

0 (jamais), 1 (presque jamais), 2 (parfois), 3 (presque toujours), ou 4 (toujours).

Concernant votre voix :

1. Ma voix semble rauque ou enrouée 0 1 2 3 4
2. J'ai l'impression que je dois forcer pour produire ma voix 0 1 2 3 4
3. Je sens une douleur ou une gêne après avoir parlé 0 1 2 3 4
4. Je sens une fatigue vocale après une longue conversation 0 1 2 3 4
5. J'ai des difficultés à produire une voix forte (par exemple : crier pour appeler quelqu'un dans la maison) 0 1 2 3 4
6. J'ai l'impression que ma voix est faible 0 1 2 3 4
7. J'ai des difficultés pour chanter 0 1 2 3 4
8. J'ai des difficultés à produire un son aigu 0 1 2 3 4
9. J'ai l'impression que le son de ma voix est grave 0 1 2 3 4
10. Je suis à bout de souffle quand je parle 0 1 2 3 4

Score relatif à la voix :

Concernant votre gorge et votre cou :

1. J'ai beaucoup de salive/mucus/glaïres dans la gorge 0 1 2 3 4
2. J'ai l'impression d'avoir quelque chose de coincé dans la gorge 0 1 2 3 4
3. Je me racle souvent la gorge car je sens beaucoup de salive/mucus/glaïres 0 1 2 3 4
4. Je sens une gêne lorsque j'avale des aliments ou des liquides 0 1 2 3 4
5. Je tousse après avoir mangé ou lorsque je suis allongé.e 0 1 2 3 4
6. J'ai la bouche ou la gorge sèche, et j'ai soif 0 1 2 3 4
7. J'ai des difficultés respiratoires ou j'ai fréquemment des épisodes d'étouffement 0 1 2 3 4
8. Mon cou est engourdi et je sens une gêne (ou une douleur) 0 1 2 3 4
9. Je sens un poids, une gêne, ou une douleur dans la poitrine 0 1 2 3 4
10. Mes épaules sont engourdies et je sens une gêne (ou une douleur) 0 1 2 3 4

Score relatif à la gorge et au cou :

Score total :

References

1. Kumar AN, Dent PC. Thyroidectomy. *Surgery*. 2024;42:246–251.
2. Rossi L, Materazzi G, Bakkar S, et al. Recent trends in surgical approach to thyroid cancer. *Front Endocrinol*. 2021;12:699805. <https://doi.org/10.3389/fendo.2021.699805>.
3. Fassari A, Gurrado A, Iossa A, et al. Definition of learning curve for thyroidectomy: systematic review on the different approaches. *Gland Surg*. 2023;12:989–1006.
4. Fortuny JV, Guigard S, Karenovics W, et al. Surgery of the thyroid: recent developments and perspective. *Swiss Med Wkly*. 2015;145:w14144.
5. INAMI - Cellule Soins Efficaces. Variations de pratiques médicales: Thyroïdectomie; 2023.
6. Wojtczak B, Sępek M, Sutkowski K, et al. Changes in thyroid surgery over last 25 years. *Sci Rep*. 2025;15:14432.
7. Tarallo M, Carlino C, Crocetti D, et al. Safety and outcomes of thyroid surgery: a high-volume center case series and the role of intraoperative neuromonitoring in preventing complications. *J Clin Med*. 2025;14:6077.
8. Gerardi I, Verro B, Amodei R, et al. Thyroidectomy and its complications: a comprehensive analysis. *Biomedicines*. 2025;13:433.
9. Ren T, Manning S, Lee JC, et al. Delayed-onset vocal cord palsy after thyroidectomy occurring despite normal initial post-operative endoscopy. *ANZ J Surg*. 2024;94:1794–1799.
10. Hayward NJ, Grodski S, Yeung M, et al. Recurrent laryngeal nerve injury in thyroid surgery: a review. *ANZ J Surg*. 2013;83:15–21.
11. Hisham AN, Lukman MR. Recurrent laryngeal nerve in thyroid surgery: a critical appraisal. *ANZ J Surg*. 2002;72:887–889.
12. Chiang F, Lu I, Chen H, et al. Anatomical variations of recurrent laryngeal nerve during thyroid surgery: how to identify and handle the

- variations with intraoperative neuromonitoring. *Kaohsiung J Med Sci.* 2010;26:575–583.
13. Potenza AS, Araujo Filho VJF, Cernea CR. Injury of the external branch of the superior laryngeal nerve in thyroid surgery. *Gland Surg.* 2017;6:552–562.
 14. Rodriguez A, Hans S, Lechien JR. Post-thyroidectomy voice and swallowing disorders and association with laryngopharyngeal reflux: a scoping review. *Laryngosc Investig Otolaryngol.* 2023;8:140–149.
 15. Söber L, Lepner U, Kirsimägi Ü, et al. Voice and swallowing disorders after thyroid surgery. *J Voice.* 2024;38:1254.e11–1254.e20.
 16. Mihai R. Voice and swallowing symptoms after thyroid surgery assessed using the ThyVoice online platform. *Eur Thyroid J.* 2023;12:e230008. <https://doi.org/10.1530/ETJ-23-0008>.
 17. Scerrino G, Tudisca C, Bonventre S, et al. Swallowing disorders after thyroidectomy: what we know and where we are. A systematic review. *Int J Surg.* 2017;41:S94–S102.
 18. Aref EEM, Ahmed GAE-H, Ibrahim RA-W, et al. Vocal dysfunction following thyroid surgery: a multidimensional subjective and objective study. *Egypt J Otolaryngol.* 2024;40:106.
 19. Galluzzi F, Garavello W. Dysphagia following uncomplicated thyroidectomy: a systematic review. *Eur Arch Oto-rhino-laryngol.* 2019;276:2661–2671.
 20. Đanić-Hadžibegović A, Hergešić F, Babić E, et al. Thyroidectomy-related swallowing difficulties: review of the literature. *Acta Clin Croat.* 2020;59:38–49.
 21. Hillenbrand A, Cammerer G, Dankesreiter L, et al. Postoperative swallowing disorder after thyroid and parathyroid resection. *Pragmat Obs Res.* 2018;9:63–68.
 22. Pereira JA, Girvent M, Sancho JJ, et al. Prevalence of long-term upper aerodigestive symptoms after uncomplicated bilateral thyroidectomy. *Surgery.* 2003;133:318–322.
 23. Krekeler BN, Wendt E, Macdonald C, et al. Patient-reported dysphagia after thyroidectomy. *JAMA Otolaryngol Head Neck Surg.* 2018;144:342.
 24. Lombardi CP, Raffaelli M, D'Alatri L, et al. Voice and swallowing changes after thyroidectomy in patients without inferior laryngeal nerve injuries. *Surgery.* 2006;140:1026–1034.
 25. Scerrino G, Inviati A, Di Giovanni S, et al. Esophageal motility changes after thyroidectomy; possible associations with postoperative voice and swallowing disorders: preliminary results. *Otolaryngol Head Neck Surg.* 2013;148:926–932.
 26. Chawaka HJ, Teshome ZB. The underreported postoperative suffering after thyroid surgery: dysphagia, dysphonia, and neck pain—a cross-sectional study. *Anesthesiol Res Pract.* 2023;2023:1312980.
 27. Thorsen RT, Døssing H, Bonnema SJ, et al. The impact of post-thyroidectomy neck stretching exercises on neck discomfort, pressure symptoms, voice and quality of life: a randomized controlled trial. *World J Surg.* 2022;46:2212–2222.
 28. Vardaxi C, Tsetsos N, Koliastasi A, et al. Swallowing disorders after thyroidectomy: a systematic review and meta-analysis. *Eur Arch Oto-rhino-laryngol.* 2022;279:4213–4227.
 29. Kuhn MA, Bloom G, Myssiorek D. Patient perspectives on dysphonia after thyroidectomy for thyroid cancer. *J Voice.* 2013;27:111–114.
 30. Kletzien H, Macdonald CL, Orne J, et al. Comparison between patient-perceived voice changes and quantitative voice measures in the first postoperative year after thyroidectomy. *JAMA Otolaryngol Head Neck Surg.* 2018;144:995.
 31. Fernandes-Taylor S, Damico-Smith C, Arroyo N, et al. Multicenter development and validation of the vocal Cord Paralysis Experience (CoPE), a patient-reported outcome measure for unilateral vocal fold paralysis-specific disability. *JAMA Otolaryngol Head Neck Surg.* 2022;148:756–763.
 32. Alnefaie SM, Alotaibi MS, Alsobaie AE, et al. Assessment of patients' quality-of-life post-thyroidectomy. *Cureus.* 2024;16:e52744. <https://doi.org/10.7759/cureus.52744>.
 33. Jacobson BH, Johnson A, Grywalski C, et al. The Voice Handicap Index (VHI). *Am J Speech Lang Pathol.* 1997;6:66–70.
 34. Wilson JA, Webb A, Carding PN, et al. The Voice Symptom Scale (VoiSS) and the Vocal Handicap Index (VHI): a comparison of structure and content. *Clin Otolaryngol Allied Sci.* 2004;29:169–174.
 35. Mathieson L, Hirani SP, Epstein R, et al. Laryngeal manual therapy: a preliminary study to examine its treatment effects in the management of muscle tension dysphonia. *J Voice.* 2009;23:353–366.
 36. McHorney CA, Robbins J, Lomax K, et al. The SWAL-QOL and SWAL-CARE outcomes tool for oropharyngeal dysphagia in adults: III. Documentation of reliability and validity. *Dysphagia.* 2002;17:97–114.
 37. Dwivedi RC, Rose SS, Roe JWG, et al. Validation of the Sydney Swallow Questionnaire (SSQ) in a cohort of head and neck cancer patients. *Oral Oncol.* 2010;46:e10–e14.
 38. Belafsky PC, Mouadeb DA, Rees CJ, et al. Validity and reliability of the Eating Assessment Tool (EAT-10). *Ann Otol Rhinol Laryngol.* 2008;117:919–924.
 39. Silbergleit AK, Schultz L, Jacobson BH, et al. The Dysphagia Handicap Index: development and validation. *Dysphagia.* 2012;27:46–52.
 40. Chen AY, Frankowski R, Bishop-Leone J, et al. The development and validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer: the M. D. Anderson dysphagia inventory. *Arch Otolaryngol Head Neck Surg.* 2001;127:870–876.
 41. Woisard V, Bodin S, Puech M. The Voice Handicap Index: impact of the translation in French on the validation. *Rev Laryngol Otol Rhinol.* 2004;125:307–312.
 42. Bourque J-M, Defoy L, Batcho CS, et al. Cross-cultural adaptation and validation of the Voice Handicap Index in the Quebec French Population (VHI-QF). *J Voice.* 2020;34:811.e1–811.e6.
 43. Côté C, Fortin J, Brais B, et al. Cross-cultural adaptation of the SWAL-QOL and the Sydney Swallow Questionnaire (SSQ) into French-Canadian and preliminary assessment for their use in an oculopharyngeal muscular dystrophy (OPMD) population. *Qual Life Res.* 2022;31:293–302.
 44. Khaldoun E, Woisard V, Verin É. Validation in French of the SWAL-QOL scale in patients with oropharyngeal dysphagia. *Gastroenterol Clin Biol.* 2009;33:167–171.
 45. Audag N, Goubau C, Danse E, et al. Validation and reliability of the French Version of the Sydney Swallow Questionnaire. *Dysphagia.* 2019;34:556–566.
 46. Woisard V, Lepage B. The 'Deglutition Handicap Index' a self-administrated dysphagia-specific quality of life questionnaire: temporal reliability. *Rev Laryngol Otol Rhinol.* 2010;131:19–22.
 47. Chun B, Bae J, Chae B, et al. Early postoperative vocal function evaluation after thyroidectomy using thyroidectomy related voice questionnaire. *World J Surg.* 2012;36:2503–2508.
 48. Nam I, Bae J, Shim M, et al. The importance of preoperative laryngeal examination before thyroidectomy and the usefulness of a voice questionnaire in screening. *World J Surg.* 2012;36:303–309.
 49. Hwang Y-S, Shim M-R, Kim G-J, et al. Development and validation of the Thyroidectomy-Related Voice and Symptom Questionnaire (TVSQ). *J Voice.* 2022;36:145.e15–145.e22.
 50. Belafsky PC, Postma GN, Koufman JA. Validity and reliability of the Reflux Symptom Index (RSI). *J Voice.* 2002;16:274–277.
 51. Nanjundeswaran C, Jacobson BH, Gartner-Schmidt J, et al. Vocal Fatigue Index (VFI): development and validation. *J Voice.* 2015;29:433–440.
 52. Kim G-J, Bang J, Shin H-I, et al. Persistent subjective voice symptoms for two years after thyroidectomy. *Am J Otolaryngol.* 2023;44:103820.
 53. Abaalkhail M, Alessa M, AIKhashan R, et al. Validity and reliability of the Arabic Thyroidectomy-Related Voice and Symptom Questionnaire (A-TVSQ) for assessing post-thyroidectomy patients. *J Voice.* 2023. [https://doi.org/10.1016/j.jvoice.2023.10.001.40\(2\):521.e7-521.e12](https://doi.org/10.1016/j.jvoice.2023.10.001.40(2):521.e7-521.e12).
 54. Paspala A, Dimitriou E, Delides A, et al. Translation and cross-cultural adaptation of the thyroidectomy-related voice questionnaire (TVSQ) in Greek population. *J Voice.* 2024. [https://doi.org/10.1016/j.jvoice.2024.06.006.S0892-1997\(24\)00179-6](https://doi.org/10.1016/j.jvoice.2024.06.006.S0892-1997(24)00179-6).

55. Choubineh M, Bafetrat A, Nikpoor N, et al. Translation, cross-cultural adaptation and preliminary validation of the Thyroidectomy-Related Voice and Symptom Questionnaire (TVSQ) into Persian. *J Voice*. 2025. [https://doi.org/10.1016/j.jvoice.2025.10.042.S0892-1997\(25\)00465-5](https://doi.org/10.1016/j.jvoice.2025.10.042.S0892-1997(25)00465-5).
56. Peix J-L, Lifante J-C, Maillard L. Évolution récente de la chirurgie thyroïdienne. *Bull Acad Natl Med*. 2017;201:707–713.
57. Colonna M, Borson-Chazot F, Delafosse P, et al. Progression of incidence and estimate of net survival from papillary thyroid cancers diagnosed between 2008 and 2016 in France. *Ann Endocrinol*. 2020;81:530–538.
58. Li M, Meheus F, Polazzi S, et al. The economic cost of thyroid cancer in France and the corresponding share associated with treatment of overdiagnosed cases. *Value Health*. 2023;26:1175–1182.
59. Francart J, Van Den Bruel A, Decollonne B, et al. *Différences géographiques de l'incidence du cancer de la thyroïde en Belgique*. Centre Fédéral d'Expertise des Soins de Santé (KCE); 2012 <https://doi.org/10.57598/R177BS>.
60. Cruchinho P, López-Franco MD, Capelas ML, et al. Translation, cross-cultural adaptation, and validation of measurement instruments: a practical guideline for novice researchers. *J Multidiscip Healthc*. 2024;17:2701–2728.
61. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation Process for Patient-Reported Outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health*. 2005;8:94–104.
62. World Health Organization. WHODAS 2.0 Translation Package: Translation and Linguistic Evaluation Protocol and Supporting Material (Version 1.0). 2010.
63. Chalmers J, Armour M. The Delphi technique. *Handbook of Research Methods in Health Social Sciences*. Singapore: Springer; 2019:715–735.
64. McPherson S, Reese C, Wendler MC. Methodology update: Delphi studies. *Nurs Res*. 2018;67:404–410.
65. Hsu C-C, Sandford BA. The Delphi technique: making sense of consensus. *Pract Assess Res Eval*. 2007;12:10. <https://doi.org/10.7275/pdz9-th90>.
66. von der Gracht HA. Consensus measurement in Delphi studies. Review and implications for future quality assurance. *Technol Forecast Soc Chang*. 2012;79:1525–1536.
67. Diamond IR, Grant RC, Feldman BM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. *J Clin Epidemiol*. 2014;67:401–409.
68. The jamovi project. jamovi. 2025.
69. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53–55.
70. George D, Mallery P. *IBM SPSS Statistics 23 Step by Step*. New York: Routledge; 2016 <https://doi.org/10.4324/9781315545899>.
71. Schober P, Boer C, Schwarte LA. Correlation coefficients. *Anesth Analg*. 2018;126:1763–1768.
72. Cureton EE. Rank-biserial correlation. *Psychometrika*. 1956;21:287–290.
73. Glass GV. Note on rank biserial correlation. *Educ Psychol Meas*. 1966;26:623–631.
74. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3:77–101.
75. Influence Communication. Scolarius - mesurer pour mieux se faire comprendre; 2010 [cited 2025 Nov 17]. Available at: <https://scolarius.com/>.
76. Wittich L, Tsatsaronis C, Kuklinski D, et al. Patient-reported outcome measures as an intervention: a comprehensive overview of systematic reviews on the effects of feedback. *Value Health*. 2024;27:1436–1453.
77. Greenhalgh J, Gooding K, Gibbons E, et al. How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. *J Patient Rep Outcomes*. 2018;2:42.
78. Elsheikh E, Quriba AS, El-Anwar MW. Voice changes after late recurrent laryngeal nerve identification thyroidectomy. *J Voice*. 2016;30:762.e1–762.e9.
79. Lombardi CP, Raffaelli M, D'Alatri L, et al. Voice and swallowing changes after thyroidectomy in patients without inferior laryngeal nerve injuries. *Surgery*. 2006;140:1026–1034.
80. Banks CA, Ayers CM, Hornig JD, et al. Thyroid disease and compressive symptoms. *Laryngoscope*. 2012;122:13–16.
81. Sinagra DL, Montesinos MR, Tacchi VA, et al. Voice changes after thyroidectomy without recurrent laryngeal nerve injury. *J Am Coll Surg*. 2004;199:556–560.
82. Delgado-Vargas B, Lloris Romero-Salazar A, Cobeta I. Vocal changes following thyroid surgery: prospective study of objective and subjective parameters. *J Voice*. 2019;33:27–32.
83. Kim S-Y, Kim G-J, Lee D-H, et al. Analysis of voice changes after thyroidectomy using the thyroidectomy-related voice and symptom questionnaire. *Auris Nasus Larynx*. 2021;48:963–972.
84. Nam I, Bae J, Lee S, et al. Prospective voice assessment after uncomplicated thyroidectomy: a comprehensive analysis of a single centre experience. *Clin Otolaryngol*. 2023;48:39–49.