

**Frameworks, Terminology and Definitions Used for the
Classification of Voice Disorders: A Scoping Review**

Author

Payten, CL, Chiapello, G, Weir, KA, Madill, CJ

Published

2022

Journal Title

Journal of Voice

Version

Version of Record (VoR)

DOI

[10.1016/j.jvoice.2022.02.009](https://doi.org/10.1016/j.jvoice.2022.02.009)

Rights statement

© 2022 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/>)

Downloaded from

<http://hdl.handle.net/10072/417475>

Griffith Research Online

<https://research-repository.griffith.edu.au>

Frameworks, Terminology and Definitions Used for the Classification of Voice Disorders: A Scoping Review

*[†]Christopher L. Payten, *Greg Chiapello, *[‡]Kelly A. Weir, PhD, and [†]Catherine J. Madill, PhD, *[‡]§Southport, and [†]Camperdown, Australia

Summary: Background. A challenge for clinicians and researchers in laryngology is a lack of international consensus for an agreed framework to classify homogenous groups of voice disorders. Consistency in terminology and agreement in how conditions are classified will provide greater clarity for clinicians and researchers.

Objective. This scoping review aimed to examine the published literature on frameworks, terminology, and criteria for the classification of voice disorders.

Design. Seven online databases (MEDLINE, Embase, CINAHL, PsycInfo, Scopus, Cochrane Collaboration, Web of Science) and grey literature sources were searched. Studies published from 1940 to 2021 were included if they provided a descriptive detail of a classification framework structure and described the methodological approaches to determine classification. A narrative synthesis of the main concepts including terminology, classification criteria, grouping of conditions, critical appraisal items and gaps in research was undertaken.

Results. A total of 2,675 publications were screened. Twenty sources met inclusion criteria, including published articles and grey literature. Thirty-five classification groups and over 150 sub-groups were described. The classification group labels, and criteria for inclusion of conditions varied across the frameworks. Several key themes in terminology and criteria useful for classification are discussed, and a core set of suggested terms and definitions are presented.

Conclusions. The quality of research on classification frameworks for voice disorders is low and not one system encompasses all voice disorders across the whole spectrum. Continued high quality research using consensus methodology and inter-rater reliability scores is recommended to develop and test an internationally agreed classification framework for voice disorders.

Key Words: Voice disorder classification—Voice disorder terminology—Classification criteria—Classification framework.

INTRODUCTION

A reliable voice is an essential component of everyday communication.¹ The voice can become disordered when the quality, pitch, loudness, flexibility and vocal effort are perceived to be different from others of a similar age, gender or cultural group.² An estimated 1.7% of the general population report vocal symptoms each year, with an increasing number seeking medical referrals for their symptoms.³ Dysphonia may be a symptom of a serious medical condition including neurological impairment and head/neck cancer, requiring time-sensitive medical diagnosis and management.⁴ However, 40% to 60% of patients referred for voice assessment do not have any organic structural or

neurological pathophysiology, with the vocal disorder resulting from other *non-organic* causes.⁵ Cost-effective treatment planning and prevention of symptom decline requires early multi-disciplinary assessment and accurate diagnosis.⁶

The term “voice disorder” embodies many different conditions with signs and symptoms that can present in isolation or combination with each other.⁴ As voice production involves the complex synergy of multiple physiological, biological, and psychosocial sub-systems, aetiology of the disorder can often be multifactorial. The process of classifying the cause of vocal decline, beyond identifying visible or organic laryngeal pathophysiology, is complex, and requires the expertise of multiple professionals.⁷

The multidisciplinary team contributing to the diagnosis of a voice disorder includes a combination of Otolaryngologists (Ear, Nose and Throat, ENT), Speech-Language Pathologists (SLP), voice scientists, psychologists, singing teachers and vocal coaches. Each professional will focus on a different level of the vocal symptoms to classify the vocal complaint. For example, ENTs may describe a voice disorder based on presence or absence of visually observed pathophysiology, whereas SLPs may describe voice disorders using subjective auditory perceptual judgements. Consequently, each professional may use their own medical jargon to describe similar voice symptoms or presentations. Clinicians will frequently use broad classification terms (eg, non-specific dysphonia, chronic laryngitis, functional dysphonia,

Accepted for publication February 6, 2022.

This review was supported by the Gold Coast Hospital and Health Service Allied Health Research Clinical Backfill Scheme for the 2 reviewers (CP, GC).

From the *Department of Speech Pathology and Audiology, Gold Coast Health, Gold Coast University Hospital, Southport, Queensland, Australia; †Faculty of Medicine and Health, Sydney School of Health Sciences, Discipline of Speech Pathology, University of Sydney, Camperdown, New South Wales, Australia; ‡Department of Allied Health Research, Gold Coast Health, Gold Coast University Hospital, Southport, Queensland, Australia; and the §Menzies Health Institute Queensland, School of Health Sciences, Griffith University, Gold Coast Campus, Southport, Queensland, Australia.

Address correspondence and reprint requests to Christopher L Payten, Speech Pathology Department, Gold Coast University Hospital, 1 Hospital Boulevard, Southport, Queensland 4215, Australia. E-mail: christopher.payten@health.qld.gov.au

Journal of Voice, Vol. ■■■, No. ■■■, pp. ■■■–■■■
0892-1997

© 2022 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.2022.02.009>

or non-organic dysphonia) to define homogenous groups of voice disorders where there is no distinct criteria to determine a specific diagnosis. Conversely, specific diagnostic terminology may be consistently used across professionals where there is visible laryngeal pathophysiology, regardless of the aetiology (eg, vocal fold nodules, vocal fold polyp, vocal fold paralysis). However, for many patients, there may be no single objectively identifiable cause for their vocal symptoms, and more than one condition may be contributing to multiple diagnoses and/or descriptors.⁸

Consequently, terminology for similar groups of voice disorders in the literature is often used interchangeably and without clear definition, specifically when the diagnostic nomenclature for observed organic pathology is not well defined or fully understood.^{7,9,10} Consistent classification terminology is important when diagnostic criteria for specific diseases are not available or are not considered useful for making a diagnosis. Clinicians frequently rely on classification criteria to inform their diagnostic impression, or to plan and communicate effective treatment recommendations.¹¹ Therefore, consistency in the terminology and criteria for classification of homogenous groups of voice disorders is needed to provide greater clarity for clinicians when communicating assessment findings and treatment outcomes.⁷ Additionally, clearly defined terminology and classification criteria will better inform the design of future studies that aim to measure the diagnostic value of the assessment process and test measures for this population.¹⁰

A useful classification system is one which is reliable and valid, and typically developed using one of two approaches.¹² The statistical approach uses statistical procedures to group patients with similar characteristics to demonstrate that the different groups do not have overlapping attributes. The judgment approach may rely on three forms of judgment: i) traditional custom, where the developer identifies the most important variables represented in the literature, ii) conventional wisdom, where the developer relies on common unpublished beliefs or clinician consensus to decide which variables should be included, and iii) personal experience. When using classification systems as a diagnostic tool, it is important to consider the clinical utility of the system and whether it provides a reliable description of the underlying disorders.¹²

Review Objectives

The primary objective of this review was to systematically examine the published literature to identify and evaluate the existing classification frameworks for voice disorders. We sought to examine themes in the different terminology used; describe the types of diagnoses and their diagnostic criteria (signs, symptoms, test results) that fit within each of the classification groups; and describe the types of clinical assessments used to determine classification. Additionally, we aimed to critically appraise studies that describe classification frameworks for voice disorders according to published methodological criteria¹² in order to recommend a core set of voice disorder classification groups and classification terms with corresponding criteria. This work will provide

clarity for clinicians when classifying voice disorders that share similar diagnostic features to inform future research in assessment, classification, and treatment of these conditions.

Review questions

1. What frameworks, terminology and definitions for the classification of voice disorders are described in the literature?
2. What clinical criteria exist within these frameworks to guide clinicians in the classification of voice disorders?
3. Do the existing classification frameworks satisfy methodological guidelines for the development of classification systems?

Sub-questions

1. What specific conditions or diagnoses are described within the classification groups?
2. What clinical assessments and case history information are described to guide the classification of voice disorders?

METHODS

Protocol registration

The scoping review protocol was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analysis Protocols for Scoping Reviews (PRISMA-ScR)¹³ and the Joanna Briggs Institute (JBI) methodology for scoping reviews.¹⁴ The final protocol was peer-reviewed and registered prospectively with the JBI systematic review register and published in the JBI evidence synthesis journal.¹⁵

Information sources

The following bibliographic databases were searched from 1940 to October 2021: MEDLINE (Ovid), Embase (Elsevier), CINAHL (EBSCOhost), PsycInfo (Ovid), Scopus (Elsevier), Cochrane Database of Systematic Reviews (Cochrane Library), and Web of Science (Clarivate analytics). Grey literature searches were also conducted through Google, Google Scholar, ClinicalTrials.gov, Cochrane Central Register of Controlled Trials (CENTRAL) and ProQuest Dissertations and Theses. In addition, grey literature searches included Speech-Language Pathology and ENT peak international bodies including Speech Pathology Australia (SPA), The Royal College of Speech and Language Therapists (RCSLT), The American Speech Language Hearing Association (ASHA), Australasian Society of Otolaryngology Head and Neck Surgery (ASOHNS), American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS) and European Laryngological Society (ELS).

Search strategy

A three-step search strategy was used to identify relevant sources for the review. An initial search through PubMed, textbooks, international websites, and key review papers included the broad classification terms (organic voice disorder, functional voice disorder, psychogenic voice disorder). Key terms were then determined through discussion between two authors (CP, CM) and refined using key concepts addressing classification. The search strategy was reviewed by a senior medical librarian at the University of Sydney. The final search strategy for MEDLINE can be found in the appendix. Searches were modified for each database.

The comprehensive search of all databases was conducted by the first author (CP) and a senior medical librarian at Gold Coast University Hospital with expertise in systematic review searching, on July 2, 2020. The first author conducted a final search to include new articles published to 14 October 2021. The search results were exported to EndNote X8 and the duplicates removed by the first author (CP) using the Bramer method.¹⁶

In the final step of the search strategy, reference lists of all studies identified for full text review were examined for

additional relevant sources not found during the bibliographic database search.

Eligibility criteria

Studies were included if they described a named classification framework for voice disorders. There was no age limit for participants in included studies because the same terminology is used for adults and children with voice disorders. Detailed inclusion and exclusion are described in Table 1. The year 1940 was selected because voice disorders of *functional* origin were first described around this time, and the aim of this review was to capture as wide a range of the literature as possible. Disorders of cough and laryngeal airway (ie, inducible laryngeal obstruction, paradoxical vocal fold movement) were not included in this review.

Selection of sources of evidence

After deduplication, references were uploaded to Covidence for screening.¹⁷ Titles and abstracts were independently screened for eligibility by two authors (CP, GC). When screening was unable to determine if the title or abstract met inclusion criteria, the article was selected for full text review.

TABLE 1.
Inclusion/Exclusion Criteria

	Inclusion	Exclusion
Population	Participants (children or adult) diagnosed with a voice disorder using any diagnostic criteria. Quantitative, qualitative, and mixed methods studies; systematic reviews; reports; and grey literature that aim to evaluate patients with a voice disorder and including: A) The development or description of a voice disorder classification framework and/or classification criteria (ie, groups of conditions). and B) A description of defined criteria used to describe one or more voice disorder classification group(s)/framework.	Studies including animal models.
Concept	Description of voice disorder classification or criteria or specific voice disorder diagnoses/criteria. Well-defined methodological approaches to determine a set of diagnostic criteria (eg, a consensus approach based on expert opinion; statistical modelling; systematic literature reviews; methods considering multiple diagnostic tests).	Classification systems for laryngeal airway diagnoses (ie, paradoxical vocal fold movement, inducible laryngeal obstruction), chronic cough, alaryngeal speech, elective mutism, or resonance disorders.
Context	Geographical and clinical contexts (specifically the treating professions) where a classification or diagnosis of voice disorder has been employed Articles published in English Language Articles published from 1940 to October 2021	Studies in languages other than English.

TABLE 2.
Glossary of Terms Used in the Reporting of This Review

Classification framework	A structured framework consisting of single or multiple groups of conditions.
Classification group	An overarching group where authors have described homogenous cohorts of voice disorders or sub-groups of voice disorders.
Multi-axis framework	A classification framework where authors include more than one classification group of voice disorders.
Single-axis framework	A classification framework where authors describe a single classification group.
Sub-group	A sub-group of voice disorders contained within a single classification group.
Groups of conditions	Groups of voice disorders listed within a sub-group, that is, hypo-functional neurological conditions, with a neurological subgroup.
Disorder/diagnosis/condition	A specific voice disorder contained within a classification group or within a subgroup
Analysis domain	A temporary domain group label used by the authors of this scoping review to assist with data synthesis and reporting of the terminology.

Full text reviewing was conducted independently by the same authors (CP, GC) and any conflicts were resolved through discussion. When agreement on inclusion could not be met, consensus was achieved using a third reviewer (CM).

A purposive data extraction form was developed through discussion with the research team to determine the variables to extract. Two reviewers (CP, GC) independently read and extracted the data, whilst continually discussing the results and clarifying questions during the extraction process. Extracted data included specific details about the study population (authors, year of publication, participant details); concept (classification frameworks, terminology used, classification criteria used or tested, clinical assessment data used for classification, level of evidence); context (country of study, author profession).

Data extraction and synthesis

All classification frameworks and the disorders listed were placed into a table ([supplementary file](#)). The frameworks were described in terms of overall structure and themes in terminology and classification criteria. We define *structure* as the distinct conceptualisation and differentiation of which disorders were included in the system and what criteria was used to group conditions within the framework. We attempted to reduce bias in our reporting of the extracted data by using a standardised set of terminology in this review. A glossary of terms can be found in [Table 2](#). To address the heterogenous semantic labels for the same or similar disorders described across the various frameworks, during the review process we grouped each disorder into one of three analysis domains and applied a group label (*hyperfunction-muscle tension, psychosocial, organic*) to facilitate subsequent appraisal. After reviewing all the papers and extracting the data both reviewers (CP, GC) identified the same three broad categories prior to any discussion. The analysis domain labels were selected from common terms used in the literature, and terms we concluded were the most descriptive of the entire group of disorders

contained within each analysis domain. This allowed a comparison of the terminology used by authors for similar clinical entities across frameworks.

To explore the classification criteria described by authors in their frameworks, we extracted data including details of the assessment tools described, assessment protocols, and clinical observations described by the authors. This data was included in the data extraction table ([supplementary file](#)).

Evaluation of study quality and framework development methodology

The quality of all eligible studies was assessed using relevant subsections from the JBI critical appraisal tools.¹⁴ Each study was blindly appraised by two independent reviewers (CP, GC), ratings were compared and discussed, and final consensus ratings determined. To evaluate how the classification frameworks satisfy recommended guidelines we used a previously published critical appraisal and risk-of-bias tool validated for the appraisal of classification systems, to evaluate a sub-set of articles (n = 7) which described multi-axis frameworks.¹² The same two reviewers (CP, GC) critically appraised and scored each study in this sub-set blindly according to seven domains: appropriateness for purpose, content validity, face validity, feasibility, construct validity, reliability, and generalizability. An allocated weighted score was agreed for each article by the same reviewers through a consensus.

RESULTS

Identified studies

The database search found a total of 2,675 publications for title and abstract review, documented in the PRISMA flow diagram ([Figure 1](#)). Independent review was conducted on 113 full-text articles and nine sources of grey literature (book chapters). A total of 20 sources met inclusion criteria and were selected for data extraction.

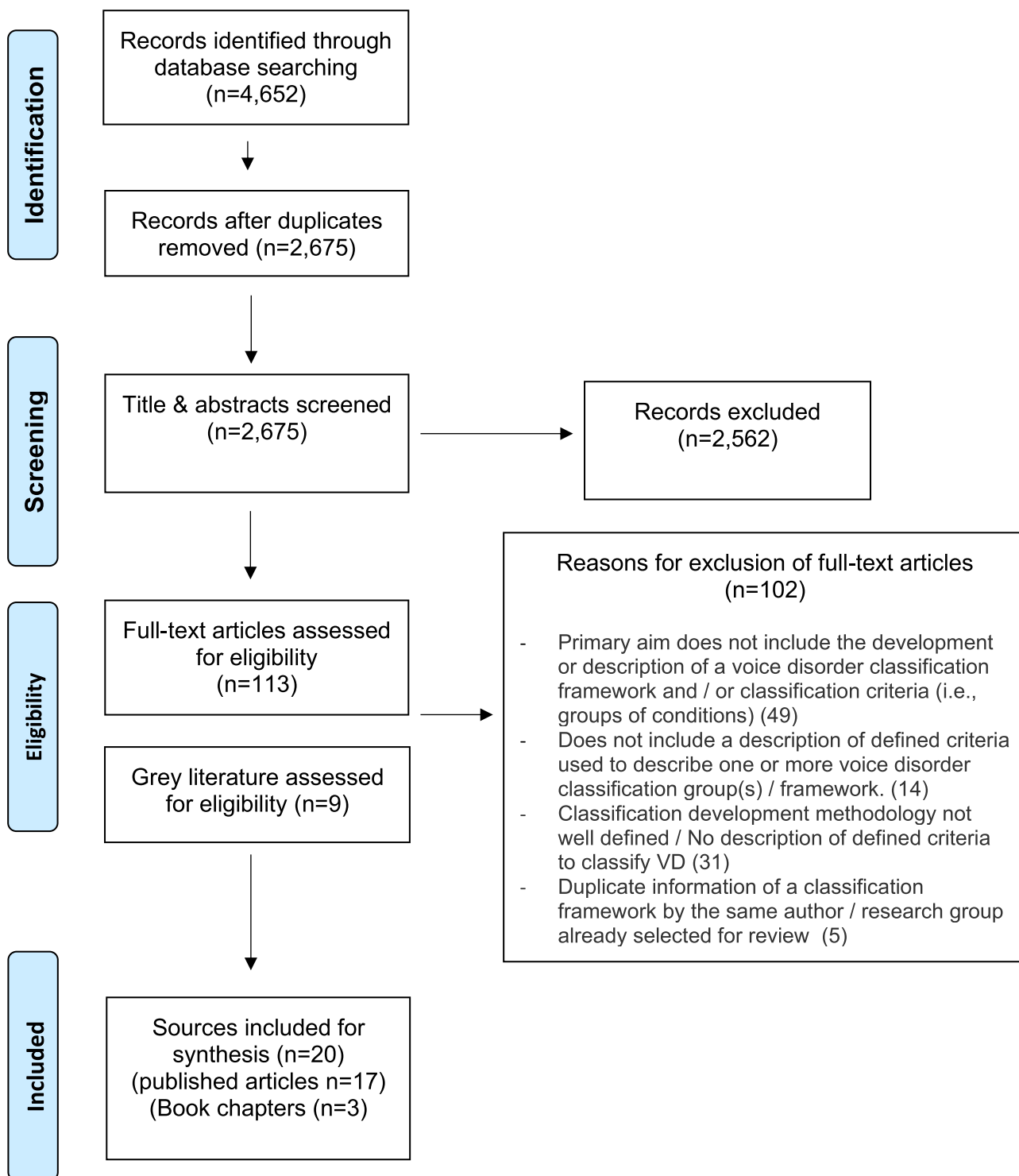


FIGURE 1. PRISMA flow diagram.

Study characteristics

Tables 3 and 4 outline the voice disorder classification labels from the 20 selected articles. A complete summary of the extracted data from each source is provided in the [supplementary file](#). The earliest paper was published in 1960, eight were published between 1960 and 2000, and

the remaining 11 were published between 2000 and 2020. Most of the studies originated from the United States ($n = 10$), followed by Canada ($n = 2$), and United Kingdom ($n = 2$). Authors from Australia, Belgium, Spain, Serbia, Netherlands, and Germany each contributed one source of evidence included in this review. The

TABLE 3A.
Multi-axis Frameworks: Classification Group Terminology and Included Subgroups/Conditions Where Classification is by Aetiology of the Voice Disorder

Morrison et al, 1986. ¹⁸ Canada, ENT/SLP/Psych	Baker et al, 2007. ⁷ Australia, SLP/ENT
<p>Functional (psychological) Dysphonia (FD). Hyper adducting ventricular band dysphonia. FD with bowed vocal folds. FD with hyperadduction (hysterical aphonia). FD with non-specific features.</p> <p>Muscular Tension Dysphonia (MTD). MTD 1 (posterior chink). MTD 2a (vocal nodules). MTD 2B (chronic laryngitis). MTD 2C (polypoidal degeneration).</p> <p>Spasmodic dysphonia. Voice disorder which relates to both organic and functional disturbances.</p>	<p>Functional Psychogenic Voice Disorder (PVD). PVD Type 1 (aphonia). PVD Type 2 (dysphonia). PVD Type 3 (Psychogenic spasmodic dysphonia). PVD Type 4 (Puberphonia/mutational falsetto).</p> <p>Functional Muscle Tension Voice Disorder (MTVD). Type 1 (no secondary pathology). Type 2a (secondary pathology). Type 2b (with secondary pathology). Type 2c (with secondary pathology).</p> <p>Organic Voice Disorder (OVD) Type 1 (mass lesions of tissue changes). Type 2 (laryngeal trauma). Type 3 (neurological LMN). Type 4 (neurological Adductor/abductor spasmodic dysphonia). Type 5 (neurological UMN with dysarthrophonia).</p>

Abbreviations: ENT, Ear Nose and Throat Surgeon; LMN, lower motor neurone; Psych, Psychiatrist; SLP, Speech Language Pathologist; UMN, upper motor neurone; VF, vocal fold.

predominant author profession was ENT (n = 8), followed by combined ENT and SLP (n = 5), then SLP and speech scientists (n = 2), and SLP and psychologists (n = 2). Seven of the 20 sources described multi-axis frameworks, and thirteen detailed a single axis framework.

Data synthesis

Comparisons of classification framework structure

The multi-axis frameworks all included a comprehensive structure with a range of disorders represented and mostly grouped according to non-organic and organic characteristics (Table 3a, Table 3b and Table 3c). Four of the seven authors divided their framework into three classification groups.^{7,18-20} Whereas Miltutinovic included two,²¹ Bradley four,⁸ and Verdolini and colleagues of the American Speech-Language-Hearing Association Special Interest Division 3 (ASHA SIG-3) classification manual for voice disorders-I described nine classification groups.⁹ There were similarities in the characteristics of voice disorders represented across the frameworks. However differences in the terminology used to label the classification groups, subgroups, and conditions represented within them. Furthermore, patterns in the methods used by authors in how they chose to group conditions within the multi-axis frameworks existed. Two authors grouped conditions according to the primary aetiology and the observed features of the voice disorder (Table 3a).^{7,18} Several authors grouped and labelled conditions based on the presence or absence of organic pathology (Table 3b).¹⁹⁻²¹ Whereas two authors classified conditions based on the location and characteristics of the

primary aetiology contributing to the voice disorder, for example, *inflammatory conditions, structural changes, neuromuscular and muscle tension imbalances* (Table 3c).^{8,9}

All but three of the thirteen single-axis frameworks described conditions of non-organic characteristics under a single classification group (Table 4a).²²⁻³¹ Arnold included a range of disorders of both non-organic and organic aetiologies grouped together under a single group label *ventricular dysphonia* (Table 4a).³⁴ Whereas two articles by Rosen and colleagues described frameworks for specific disorders, namely *vocal fold motion impairment* resulting from neurogenic, mechanical and pathological causes and *benign vocal fold lesions (BVFL)* (Table 4b).^{32,33}

Comparing classification group terminology and included conditions

Thirty-five classification groups were described across the sources, and more than 150 subgroups or conditions. The most frequently used labels for the classification groups included: *muscle tension, muscular tension*,^{7,8,18,26-28} *functional*,^{7,18,20-23} *psychogenic, psychological*,^{9,30} and *organic*.^{7,19-21} Despite the use of similar terminology, there were differences in how subgroups or conditions were placed under these labels, specifically where *muscle tension* and *functional* were concerned. For example, some authors included conditions pertaining to psychosocial and muscle tension characteristics under the classification group label *functional* (Table 3b and Table 4a),^{20,21,23} whereas others placed these conditions under *muscle tension* (Table 3a and Table 4a).^{8,25,28} Only three authors used the term *functional* exclusively to represent voice disorders of psychogenic aetiology.^{7,18,22} In our

TABLE 3B.

Multi-axis Frameworks: Classification Group Terminology and Included Subgroups/Conditions Where Classification is by the Presence or Absence of Organic Pathology

Damste, 1973. ²⁰ Netherlands, SLP	Milutinovic, 1996. ²¹ Serbia, ENT	Rosen and Murray, 2000. ¹⁹ USA, ENT/SLP
Functional dysphonia. Psychogenic aphonia. Psychogenic dysphonia. Dysphonia spastica. Habitual dysphonia. Habitual dysfunction during/after mutation.	Functional voice disorder. Phononeurosis (<i>dysphonia arising from psychogenic superimposition</i>). Phonoponosis (<i>dysphonia arising from disharmonic activity</i>).	Nonorganic voice disorder. Muscle Tension Aphonia (MTA) (<i>psychogenic voice disorder</i>). Primary muscle tension dysphonia (MTD).
Secondary organic dysphonia (<i>consequences of vocal strain</i>). Simple laryngitis. Chronic nodular laryngitis. Vocal fold oedema. Polyps. Chronic hypertrophic laryngitis (contact ulcer, leucoplakia).	Organic voice disorder. Disorders of feedback mechanisms of voice control. Congenital disorders. Inflammatory processes. Chronic mucosal irritation. Respiratory. Hormonal disorders.	Organic voice disorder. Epithelium changes. Lesions of the lamina propria. Arytenoid disorders. Other.
Primary organic dysphonia Congenital malformation. Paralysis. Changes by endocrine disorder. Specific infections. Trauma. Tumours.	Myopathy of the laryngeal muscles. Disorders of the nervous system. Syndrome of excessively inappropriate pitch (IVP). Trauma. Benign growths. Malignant tumours.	Movement disorders of the larynx. VF paralysis and other pathological conditions of the CA joint. VF paresis/atrophy. Other laryngeal movement disorders.

Abbreviations: CA, crico-arytenoid; ENT, Ear Nose and Throat Surgeon; SLP, Speech Language Pathologist; USA, United States of America; VF, vocal fold.

comparison, we could not find an association across the sources to explain these differences, including the year of publication, country of author(s), or professional group.

Classification of trauma induced mucosal changes (eg, vocal fold polyps, vocal fold nodules) was also a key difference across the frameworks. Most frameworks included mucosal changes as a subgroup under a *muscle tension* classification.^{7-9,18} In these frameworks' authors used the prefix or suffix *primary*, to denote muscle tension with no laryngeal pathology, (ie, *primary muscle tension dysphonia*, *primary muscle tension imbalance*, *muscle tension dysphonia-primary*) or *secondary* to define a compensatory laryngeal muscle tension in response to laryngeal pathology.^{7-9,27} Furthermore, the ASHA SIG-3 group advocate the term *muscle tension/adaptive dysphonia (secondary)* to reflect the compensatory maladaptation of laryngeal patterns in response to a primary organic cause.⁹ In support of this, Hillman and colleagues introduced a theoretical model of *vocal hyperfunction* and described this as "increased and less well-regulated laryngeal muscle tensions" [Hillman et al, 1989, p.374]. They report impaired muscle tensioning can result in varying risks for vocal fold collision forces and subsequent vocal fold trauma.²⁴ On the other hand, according

to this framework, increased stiffness of the vocal folds from maladaptive tensioning, in the absence of organic pathology, impairs glottic closure during phonation.²⁴ The same authors updated their terminology to *nonphonotraumatic vocal hyperfunction (NVH)* and *phonotraumatic vocal hyperfunction (PVH)* in 2020.²⁵ Conversely, other authors appeared to exclude mucosal changes resulting from trauma from their respective non-organic classification groups. For example, Rosen described *lesions of the epithelium, lamina propria* and *arytenoid*, and Milutinovic grouped *benign growths*, under their *organic voice disorder* classification groups.^{19,21} Similarly, Damste placed these conditions into a group labelled *secondary organic*.²⁰

With respect to voice disorders commonly referred to as *organic* in the literature, most frameworks included a broad list of conditions under a single classification group label (eg, organic, primary organic).^{7,19-21} However, both Bradley and the ASHA SIG-3 group described multiple classification groups, where conditions were classified according to their broad similarities in underlying aetiology.^{8,9} Examples of these classification group labels include *structural, inflammatory conditions, trauma or injury, neuromuscular or movement disorders*.

TABLE 3C.
Multi-axis Frameworks: Classification Group Terminology and Included Subgroups/Conditions Where Classification is by the Location and Characteristics of Aetiology

Verdolini et al, 2005. ⁹ USA, SLP/ENT	Bradley 2010. ⁸ UK, ENT
<p>Inflammatory conditions of the larynx. Cricoarytenoid and cricothyroid arthritis. Acute laryngitis. Laryngopharyngeal reflux. Chemical sensitivity or irritable larynx syndrome.</p>	<p>Inflammation. Infective. <i>(Primary laryngeal: viral/bacterial, fungal, secondary: pulmonary/rhinosinusitis).</i> Non infective. <i>(LPR, allergy, trauma/irritation, autoimmune, nonspecific).</i></p>
<p>Systemic conditions affecting voice. Endocrine. Immunologic. Musculoskeletal conditions affecting voice.</p>	
<p>Non laryngeal aerodigestive disorders affecting voice. Respiratory diseases affecting voice. Digestive. Infectious diseases of the aerodigestive tract. Mycotic (fungal) infections.</p>	
<p>Structural pathologies of the larynx. Malignant laryngeal lesions. Epithelial and lamina propria abnormalities of the vocal fold. Vascular anomalies of the vocal fold. Congenital and maturational changes affecting the vocal fold.</p>	<p>Neoplastic/Structural. Benign deficits/tethering. Microvascular lesions. Endocrinological. Inflammatory mass. Laryngeal framework trauma. Neoplasms. Laryngoceles. Mixed/reactive. Malignant/premalignant. Minor salivary gland.</p>
<p>Trauma or injury of the larynx. Internal laryngeal trauma. External laryngeal trauma.</p>	
<p>Neurological disorders affecting voice. Peripheral nervous system pathology. Movement disorders affecting the larynx. Central nervous system disturbance.</p>	<p>Neuromuscular. Hypo-functional. Hyper-functional. Mixed or variable hypo-/hyper-functional. Control/coordination.</p>
<p>Other disorders affecting voice. Muscle tension dysphonia (Primary). Muscle tension/adaptive dysphonia (secondary). Ventricular dysphonia. Paradoxical vocal fold movement disorder.</p>	<p>Muscle tension imbalance. Primary muscle tension imbalance. <i>(Vocal demand/strain/psychogenic/puberphonia).</i> Secondary muscle tension imbalance. <i>(Dysphonia in the presence of an underlying organic condition).</i></p>
<p>Psychiatric and psychological disorders affecting voice. Somatoform disorders. Factitious disorder. Selective mutism. Anxiety. Mood disorders. Gender identity disorder. Psychogenic polydipsia. Psychogenic tremor-like voice fluctuations.</p>	
<p>Voice disorders: undiagnosed or not otherwise specified.</p>	

Abbreviations: ENT, Ear Nose and Throat Surgeon; LPR, laryngo pharyngeal reflux; SLP, Speech Language Pathologist; UK, United Kingdom; USA, United States of America.

TABLE 4A. Single-axis Frameworks: Classification Group Terminology and Included Subgroups/Conditions of Non-organic/Mixed Organic and Non-organic Aetiology

	Aronson et al, 1966 ²² USA SLP/Psych ENT	Koufman and Bialock, 1982 ²³ USA ENT	Hillman et al, 1989 ²⁴ Canada ENT/SLP/Psych	Morrison and Rammage, 1993 ²⁶ Canada ENT/SLP	Butcher et al, 2007 ²⁹ UK Psychol/SLP	Van Houtte et al, 2011 ²⁷ Belgium ENT/SLP	Spencer, 2015 ²⁸ USA ENT	Hillman et al, 2020 ²⁵ USA SLP/SpSc	Fernandez et al, 2020 ²⁵ Spain ENT	Hackl et al, 2022 ³¹ Europe ENT
Ventricular Dysphonia (VD)	Functional Dysphonia.	Functional Dysphonia.	Vocal Hyperfunction (VH).	Muscle Misuse Voice Disorder.	Psychogenic Voice Disorder.	Muscle Tension Dysphonia (MTD).	Muscle Tension Dysphonia (MTD).	Vocal Hyperfunction (VH)	Muscle Tension Dysphonia (MTD)	Malregulative Dysphonia
Habitual VD.	Muteness.	Hysterical aphonia/dysphonia.	Adducted VH.	Laryngeal isometric. Lateral hyper-adduction; glottis/supraglottis.	Type 1: classical (hysterical) conversion. Type 2: cognitive behavioral conversion. Type 3: Habitual conversion.	Primary MTD. Secondary MTD.	Glottal fry. Excessively lowered speaking pitch. Register flipping. Functional aphonia. Post surgical functional dysphonia.	Phonotraumatic VH. Non Phonotraumatic VH.	Grade1 A-P compression. Grade2 A-P compression. Grade3 A-P compression. Grade1 Lat compression. Grade2 Lat compression. Grade3 Lat compression.	No sub-groups included
Emotional VD.	Continually whispered speech.	Habituated hoarseness.	Non adducted VH.	Supraglottic A-P contraction.						
Paralytic VD.	Intermittently phonated	Falsetto.		Conversion aphonia.						
Cerebellar VD.	whispered speech.	Voice abuse.		Psychogenic bowing.						
Vicarious VD.	Continually phonated speech.	Postoperative dysphonia.		Adolescent transitional dysphonia.						

Abbreviations: A-P, anterior posterior; ENT, Ear Nose and Throat Surgeon; Lat, lateral; Psych, Psychiatrist; Psychol, Psychologist; SLP, Speech Language Pathologist; SpSc, Speech Scientist; UK, United Kingdom; USA, United States of America; VF, vocal fold. ¹ manuscript in press at the time of review.

Comparing classification definitions and criteria

To facilitate charting and comparison of the classification criteria we allocated each classification group and/or subgroup to one of three broad analysis domains: *hyperfunction-muscle tension*, *psychosocial*, and *organic*. These provisional domain labels were selected based on the pattern of how the multi-axis frameworks are commonly structured, and terminology we felt broadly best described the conditions within each domain, as discussed in the methodology, section 2.6. Data extraction and charting of this information is captured in the [supplementary file](#). The broad themes are discussed below.

Hyperfunction—muscle tension analysis domain

The most frequent definitions used for voice disorders placed in the hyperfunction-muscle tension analysis domain included: visible and palpable muscle tension in the larynx^{7,9,19,21,22,25,26}; malregulated, maladaptive, and inefficient laryngeal tensioning,^{8,9,24,25,31} normal laryngeal condition/structure^{23,34}; normal vocal fold morphology¹⁹; absence of identifiable physical or neurological disease in the larynx^{9,21}; and demonstrated patterns of vocal fold abuse.²³ The assorted frameworks further sub-divided these groups according to various criteria, as described below.

Clinical assessment. Visible laryngeal tension of the intrinsic musculature observed with laryngoscopy was described by all authors as a defining feature of this classification group. Some further graded subgroups according to the degree of observed supraglottic/glottic constriction. Morrison described three subgroups including *isometric*, *lateral* and *anterior-posterior* glottic/supraglottic contraction.²⁶ Similarly, Fernandez defined sub-classes of Muscle Tension Dysphonia by the degree of supraglottic anterior-posterior (AP) constriction (AP1–AP3 compression).²⁹ However, supraglottic tension does not appear to be an exclusive clinical sign for these voice disorders. This is demonstrated in Arnold’s model of ventricular dysphonia, defined by faulty participation of the ventricular folds arising from multiple aetiologies, including compensatory mechanism for neurogenic impairments, or altered laryngeal structure.³⁴ Hillman and Colleagues presented a range of aetiological factors they considered important for classification to their NVH and PVH groups of *vocal hyperfunction*.²⁵ These included vocal load, palpable extrinsic laryngeal tension with phonation, changes in sub-glottic air pressures and degree of collision forces of the vocal folds. Other authors have described similar features for their *muscle tension* classification groups further supporting them as unique to this group of disorders.^{7,9,18}

Clinical features. In the frameworks where classification criteria were most clearly defined, specific history features including symptom onset, progression, and secondary symptoms were considered important when differentiating voice disorders of *hyperfunction-muscle tension* and those of *psychosocial* domains. Onset of the voice disorder in this group was frequently cited as gradual, long-standing, persistent, and fluctuating to vocal or emotional demands.^{7-9,20,23} Furthermore, secondary symptoms of vocal fatigue,

TABLE 4B.
Single-axis Frameworks: Classification Group Terminology and Included Subgroups/Conditions. Conditions of Organic Aetiology

Rosen et al, 2012 ³³	Rosen et al, 2016 ³²
USA	USA
ENT	ENT
Benign Mid-membranous Vocal Fold Lesions (BVFL)	Vocal Fold (VF) Motion Impairment
VF nodules.	VF immobility.
VF polyp.	VF hypomobility.
VF cyst.	VF paralysis.
Fibrous mass.	VF paresis.
Reactive VFL.	VF immobility/hypomobility associated with mechanical impairment of the CA joint.
Pseudocyst.	
Nonspecific VFL.	VF immobility/hypomobility related to laryngeal malignant disease.

Abbreviations: CA, cricoarytenoid; ENT, Ear Nose and Throat Surgeon; USA, United States of America.

perceived vocal effort or strain, and pain on phonation were also described exclusively in the criteria of these voice disorders. A critical point emphasised by Baker for classification to their *muscle tension voice disorder* group is the individual's ability for voluntary control over vocal production, and noticeable improvement elicited through reflex activities (ie, yawning) or tasks designed to modify the technique of phonation.⁷

Psychosocial analysis domain

Seven papers described independent classification groups that were placed into the psychosocial aetiology analysis domain (supplementary file); two sources were single-axis,^{35,36} and five were from multi-axis frameworks.^{7,9,18,20,21} Additionally, five authors included subgroups and conditions which appeared to fit to this domain but were listed as part of a heterogenous classification group overlapping with other disorders.^{20,22,23,34} Most authors defined this group as dysphonia or aphonia in the absence of laryngeal pathophysiology and in the presence of psychological or psychiatric disturbances. Classification criteria varied across the studies; however common themes were present.

Clinical assessment. Several authors defined these conditions by various laryngeal postures as observed on laryngoscopy, similarly to the *hyperfunction-muscle tension* groups. For example, Arnold defined *emotional ventricular dysphonia* as ventricular fold approximation in response to stressful episodes.³⁴ Morrison defined Types 4 - 6 in his model by the vocal fold closure pattern or absence of glottic closure during phonation, in response to psychological or emotional conflict.²⁶ However, as previously discussed, these observations appeared to overlap with other disorders that fit with conditions in other classification groups.

Auditory-perceptual assessment of vocal quality and amount of perceived voice frequently appeared in the descriptions for this group of voice disorders. Aronson and colleagues described patients diagnosed with *hysterical/conversion dysphonia*²² and grouped them into subgroups of *muteness*, *continually whispered speech*, *intermittently phonated whispered speech* and *continually phonated speech*.²²

Others described specific perceptual features of the voice under this classification, including *high-pitched*, *whispered* vocal quality, and *pitch locked*.^{7,20,23} However, the remaining sources described mixed or variable perceptual features in this group and reported similar vocal features in other classification groups of voice disorders.

Clinical features. Butcher and colleagues stressed diagnosis should ideally include psychological evaluation and, in their framework, a *psychogenic voice disorder* should not be a diagnosis of exclusion. Classification under this model was according to how well patients met the criteria for a conversion of psychological stress into physical symptoms, as defined by the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV) criteria.³⁰ Considering a slightly different aspect, the ASHA SIG-3 group categorised conditions in this classification group according to the various terms used in the DSM-IV, where they assert the voice characteristics are a secondary response to the primary cause, being any of the identified psychiatric or psychological conditions.⁹ In contrast to other frameworks, where a voice disorder is present and appears to have a psychogenic component but does not meet DSM-IV criteria for a psychiatric or psychological disorder, in the ASHA SIG-3 model it is categorised as *Undiagnosed or not otherwise specified*.⁹

Baker and colleagues described four types of *psychogenic voice disorder* in their Diagnostic Classification System for Voice Disorders (DCSVD).⁷ In addition to the detailed summary of defining features of this classification group and each subgroup, they included stressful life events or experiences of trauma may co-occur with a sudden onset of vocal symptoms. However, the authors also assert that many patients may not always describe psychosocial experiences at the onset of symptoms.⁷ The number one defining feature for the *psychogenic voice disorder* classification group in Baker's framework is a loss of volitional control over the initiation and maintenance of normal voice without neurological or structural pathology.⁷ In agreement with this, many authors stated improved phonation in this group of patients is not easily influenced by vocal technique,

despite normal laryngeal behaviors or a normally perceived voice observed during incidental vegetative tasks such as coughing or laughing.^{7,22,30}

Organic analysis domain. The nomenclature used to describe voice disorders we placed into the organic analysis domain for this review was sourced from 10 papers, of which 6 were multi-axis frameworks.^{7-9,19-21} Overall, 21 classification groups were described and 94 subgroups or conditions (Tables 3 and 4). Many frameworks did not provide sufficient detail of the classification criteria for all subgroups/conditions within their frameworks (supplementary file). As such, we could only extract data from 68% of the sources for this domain. Most frameworks grouped conditions according to broad themes of underlying structural, mucosal or neurogenic pathophysiology. These range from three classification groups encompassing a range of conditions with shared features,^{7,19,20} to 12 separate classification groups.⁹ The themes for how these were organised across the frameworks are discussed below.

Vocal fold pathology/inflammatory conditions were included by most authors using the same or similar terminology. Certain authors grouped these conditions into a single subgroup and further defined the specific conditions contained within the group. For example, Baker’s *mass lesions or tissue changes* subgroup.⁷ Other authors split the same conditions across multiple classification groups, for example the ASHA SIG-3 *structural pathologies* and *inflammatory conditions* groups.⁹ Similarly, Bradley’s model divided these conditions into *inflammation (infective and non-infective)* and *structural/neoplastic*.⁸

Laryngeal trauma was included in most frameworks as a separate condition under a single *organic* classification group, except Bradley and Rosen.^{8,19} Rosen included trauma of the larynx within the *organic voice disorder* group,¹⁹ and Bradley as part of the *structural/neoplastic* classification group.⁸ Both frameworks had the fewest number of classification groups, which may explain this decision.

Neurological and neuromuscular conditions were clearly defined in most frameworks, with common themes in the grouping of these conditions. Bradley included four subgroups of conditions, under a single classification group defined by the patients’ presenting signs or symptoms of hypo- or hyper-muscular function.⁸ Similarly, the ASHA SIG-3 group divided a neurological conditions classification group into multiple subgroups defined by site of neurological injury or breakdown (CNS, PNS), and included a separate subgroup called *movement disorders* encompassing spasmodic dysphonia and essential laryngeal tremor amongst others.⁹ Baker described three subgroups of neurological conditions, under the classification group *organic voice disorders*, with subgroups for lower motor neurone conditions, upper motor neurone conditions, and a third group encompassing spasmodic dysphonia and essential laryngeal tremor.⁷

Voice disorders not included in the analysis domains

TABLE 5. Risk of Bias Assessment of Included Studies Using the Joanna Briggs Institute (JBI) Appraisal Tool

Text and Opinion	Damste, 1973. ²⁰	Hacki et al, 2022. ³¹	Hillman et al, 2020. ²⁵	Milutinovic, 1996. ²¹	Morrison and Rammage, 1993. ²⁶	Van Houtte et al, 2011. ²⁷	Rosen et al, 2016. ³²	Rosen and Murray, 2000. ¹⁹	Arnold and Pinto, 1960. ³⁴	Bradley, 2010. ⁸	Butcher et al, 2007. ³⁰	Verdolini et al, 2005. ⁹	Spencer, 2015. ²⁸
Is the source of the opinion clearly identified?	N	Y	Y	Y	Y	Y	Y	Y	UC	Y	Y	Y	Y
Does the source of opinion have standing in the field?	UC	Y	Y	UC	Y	UC	Y	Y	UC	UC	Y	Y	UC
Are the interests of the relevant population the central focus of the opinion?	UC	Y	Y	Y	Y	Y	Y	Y	UC	Y	Y	Y	Y
Is the stated position the result of an analytical process, and is there logic in the opinion expressed?	UC	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y
Is there reference to the extant literature?	N	UC	UC	N	Y	Y	N	N	N	UC	Y	N	Y
Is any incongruence with the literature/sources logically defended?	N	UC	UC	NA	N	N	UC	N	UC	UC	Y	N	Y

Abbreviations: N, no; NA, not applicable; UC, unclear; Y, yes. †manuscript in press at the time of review.

TABLE 6.
Risk of Bias Assessment of Included Studies Using the Joanna Briggs Institute (JBI) Appraisal Tool

Analytical Cross-Sectional Studies	Fernandez, 2020. ²⁹
Were the criteria for inclusion in the sample clearly defined?	Y
Were the study subjects and the setting described in detail?	Y
Was the exposure measured in a valid and reliable way?	UC
Were objective, standard criteria used for measurement of the condition?	Y
Were confounding factors identified?	Y
Were strategies to deal with confounding factors stated?	UC
Were the outcomes measured in a valid and reliable way?	N
Was appropriate statistical analysis used?	N

Abbreviations: N, no; NA, not applicable; UC, unclear; Y, yes.

Only one condition that was included in numerous frameworks did not match any of the descriptions of the three chosen analysis domains. *Puberphonia/adolescent transitional dysphonia* was grouped under different classification groups across many of the frameworks. Bradley's model placed this condition in the *primary muscle tension* classification group⁸; Baker, Spencer and Morrison described it in the psychosocial grouping^{7,26,28}; and ASHA SIG-3 placed the condition in the *structural pathology*,

congenital and maturational changes affecting voice classification group.⁹ However, Verdolini defines this condition as the "consistent use of a high-pitched voice in a post-pubescent individual, usually male, without known organic cause" [Verdolini et al, 2005, p.83] often with secondary gains.⁹ It may therefore be argued that a more suitable category for this condition would be a disorder of psychosocial aetiology within this framework.

Grading systems for specific conditions

Two additional sources in our review described frameworks for specific conditions: *benign mid-membranous vocal fold lesions (BVFL)*³³ and *vocal fold motion impairment*.³² These well-structured single-axis frameworks by Rosen and colleagues included detailed nomenclature to define gradings of the respective conditions, together with inclusion and exclusion criteria to help determine classification. The proposed breakdown of terminology and the definitions are good examples of how a framework can facilitate clear communication of the aetiology, assessment findings and treatment planning.⁴²

Recommended clinical assessments

We aimed to examine themes in the clinical assessments described by authors for the purposes of classification, as charted in the [supplementary file](#). Most of the authors did not include specific details of the test measures, cut off scores or recommended minimum assessment protocols for classification. Where these were described, the most frequent assessment measures included indirect laryngoscopy (n=13), case history features (n=10), auditory perceptual voice assessment features (n=4), and neck/extrinsic

TABLE 7.
Risk of Bias Assessment of Included Studies Using the Joanna Briggs Institute (JBI) Appraisal Tool

Case Series	Baker et al, 2007. ⁷	Aronson et al, 1966. ²²	Koufman and Ballock, 1982. ²³	Morrison et al, 1986. ¹⁸	Rosen et al, 2012. ³³
Were there clear criteria for inclusion in the case series?	Y	Y	Y	UC	Y
Was the condition measured in a standard, reliable way for all participants included in the case series?	Y	N	N	Y	Y
Were valid methods used for identification of the condition for all participants included in the case series?	Y	UC	N	UC	Y
Did the case series have consecutive inclusion of participants?	Y	UC	UC	Y	UC
Did the case series have complete inclusion of participants?	Y	UC	UC	Y	N
Was there clear reporting of the demographics of the participants in the study?	Y	N	N	N	N
Was there clear reporting of clinical information of the participants?	N	N	N	N	Y
Were the outcomes or follow up results of cases clearly reported?	NA	NA	N	NA	Y
Was there clear reporting of the presenting site(s)/clinic (s) demographic information?	NA	N	N	UC	Y
Was statistical analysis appropriate?	Y	N	N	N	Y

Abbreviations: N, no; NA, not applicable; UC, unclear; Y, yes.

TABLE 8.
Risk of Bias Assessment of Included Studies Using the Joanna Briggs Institute (JBI) Appraisal Tool

Case Control Studies	Hillman et al, 1989. ²⁴
Were the groups comparable other than the presence of disease in cases or the absence of disease in controls?	UC
Were cases and controls matched appropriately?	UC
Were the same criteria used for identification of cases and controls?	UC
Was exposure measured in a standard, valid and reliable way?	No
Was exposure measured in the same way for cases and controls?	UC
Were confounding factors identified?	Y
Were strategies to deal with confounding factors stated?	Y
Were outcomes assessed in a standard, valid and reliable way for cases and controls?	UC
Was the exposure period of interest long enough to be meaningful?	NA
Was appropriate statistical analysis used?	N

Abbreviations: N, no; NA, not applicable; UC, unclear; Y, yes.

laryngeal palpation (n=3). However, it is difficult to draw comparisons or conclusions for our review based on the available evidence from the included sources.

Quality assessment

Study quality indicator ratings of all 20 papers were classified using the Joanna Briggs Institute (JBI) Level of Evidence (LOE) descriptions.¹⁴ Each paper was critically appraised using the relevant subsection of the JBI critical appraisal tool¹⁴ and are detailed in Tables 5, 6, 7 and 8. Most of the sources (n=13) were articles of text and opinion. The remaining seven papers included studies using a case series (n=5), case-controlled study (n=1) and analytical cross-sectional study (n=1), indicating the sources of evidence are generally of low quality.

Quality indicator ratings for the sub-set of multi-axis frameworks using the Buchbinder tool for all parameters, including overall quality rating (total score/24), and overall weighted score (mean score of 1 for each category, total score/7) are detailed in Table 9. Four articles had a mean weighted score of less than 3/7, indicating a low level of quality. The remaining 3 studies demonstrated a moderate quality rating with a mean weighted score of greater than 4, one study scoring 5.93. Most studies were author opinion, only 1 study included inter-rater reliability measures.⁷

DISCUSSION

This scoping review intended to present an overview of the existing frameworks for the classification of voice disorders,

including the terminology, definitions, classification criteria and the conditions or diagnoses grouped within the frameworks. In addition, we evaluated the quality of the evidence with respect to published methodological guidance for development of classification frameworks, to recommend a core set of voice disorder classification groups with corresponding classification criteria based on the available evidence. Evidence for clinical assessments and relevant clinical features used in classification was also extracted during the review process, however there was insufficient information for detailed analysis of this information.

Frameworks, terminology, and definitions for classification of voice disorders

A total of 20 sources were included in the review and have been summarised according to provisional analysis group labels of *hyperfunction-muscle tension*, *psychosocial* and *organic* aetiologies to aid synthesis of this data. All but one disorder type was easily attributed to one of these three groups. Several sources were incomplete and didn't encompass conditions across all our chosen domains for analysis; whereas others included single classification group labels suggesting one analysis domain over another, but with a heterogenous list of sub-groups or conditions that spanned multiple domains. Although the classification group and subgroup labels provided in these sources varied in frequency, some terms stood out more often than others. This perhaps suggests these represent more developed concepts and have a wider acceptability than those less frequently used.

Clinical criteria to guide clinicians in classification

Less than half of the sources provided well-defined classification criteria. Where definitions were included, they were frequently inconsistent with studies using the same or similar terms. Most of the inconsistencies in the criteria described were those placed in our *hyperfunction-muscle tension* and *psychosocial* analysis domains. This is not surprising given these disorders can have a variety of signs or symptoms that are not exclusive to a single underlying condition, and perhaps not easily distinguished through medical examination alone. This may also reflect the lack of agreed terminology and defining characteristics available in the literature to date.¹⁰ Given the variety in classification criteria provided in these sources, several common themes have been identified which can be considered important in the differential classification of these voice disorders. This includes the clinical history of the symptom progression,⁷⁻⁹ secondary signs and observations of increased intrinsic and extrinsic laryngeal tension,^{7-9,18,24-25} and volitional control of the individual to modify their voice with directed instruction from a voice professional.^{7,30} For conditions identified in our organic analysis domain, limited criteria could be extracted from the published frameworks. The most comprehensive frameworks included a wide-ranging list of organic conditions grouped in a similar pattern, according

TABLE 9.
Risk of Bias of Multi-axis Frameworks Using the Buchbinder Risk of Bias Tool

			Damste, 1973. ²⁰	Morrison et al, 1986. ¹⁸	Milutinovic, 1996. ²¹	Rosen and Murray, 2000. ¹⁹	Baker et al, 2007. ⁷	Bradley, 2010. ⁵	Verdolini et al, 2005. ⁹	
Purpose	I	Is the purpose, population, and setting clearly specified?	P	Y	P	Y	Y	Y	Y	
	I	Is the domain and all specific exclusions from the domain clearly specified?	P	P	Y	N	Y	P	Y	
		Are all relevant categories included?	Y	Y	Y	Y	Y	Y	Y	
	II	Is the breakdown of categories appropriate, considering the purpose?	Y	Y	N	P	Y	Y	Y	
		Are the categories mutually exclusive?	P	Y	P	P	Y	Y	Y	
	III	Was the method of development appropriate?	N	P	N	N	Y	N	Y	
Face validity	IV	If multiaxial, are criteria of content validity satisfied for each additional axis?	P	Y	N	N	Y	P	Y	
	I	Is the nomenclature used to label the categorised satisfactory?	P	Y	N	Y	Y	Y	Y	
		Are the terms used based upon empirical (ie, directly observable) evidence?	DK	P	N	N	Y	DK	Y	
	II	Are the criteria for determining inclusion into each category clearly specified?	N	P	N	P	Y	Y	Y	
		If yes, do these criteria appear reasonable?	NA	Y	NA	N	Y	Y	Y	
		Have the criteria been demonstrated to have validity and/or reliability?	DK	N	N	N	DK	DK	DK	
Feasibility	III	Are the definition of criteria clearly specified?	N	P	N	P	Y	N	Y	
	IV	If multiaxial, are criteria of face validity satisfied for each additional axis?	N	P	N	N	Y	P	Y	
	I	Is the classification simple to understand?	Y	Y	N	N	Y	Y	Y	
	II	Is the classification easy to perform?	DK	Y	P	N	Y	Y	Y	
	III	*Does it rely on clinical examination alone?	DK	Y	DK	DK	N	N	N	
	IV	†Are special skills, tools, and/or training required?	DK	N	DK	DK	Y	Y	Y	
Construct validity	V	‡How long does it take to perform?	NA	NA	NA	NA	NA	NA	NA	
	I	Does it discriminate between entities that are thought to be different in a way appropriate for the purpose?	P	DK	P	DK	Y	DK	Y	
	II	Does it perform satisfactorily when compared to other classification systems which classify the same domain?	DK	DK	N	DK	DK	DK	DK	
	Reliability	I	Does the classification system provide consistent results when classifying the same conditions (test – retest)?	DK	DK	DK	N	Y	DK	DK
		II	Is the intraobserver and interobserver reliability satisfactory?	DK	DK	DK	N	Y	DK	DK
	Generalisability	I	Has it been used in other studies and/or settings?	DK	Y	N	DK	Y	DK	Y
Overall score			4	15	4	5	19.5	10.5	17	
Overall weighted score (Potential max = 7, mean score out of 1 for each category)			1.65	4.40	1.29	1.62	5.93	2.67	4.86	

* Clinical examination = case history by ENT/SLP including perceptual judgement by ENT/SLP, visualisation of the larynx (laryngoscopy/mirror examination).

† Special skills/tools = video stroboscopy, electromyography (EMG), other.

‡ How long does it take to perform = not judged.

Abbreviations: DK, don't know; N, no; NA, not applicable; P, partial; Y, yes.

Scores calculated by summing number of yes (1 point) and partially (0.5 point) responses, except for the item "are special skills, tools and/or training required?" for which a "no" response = 1 point.

to their underlying aetiology with respect to the location of primary breakdown (eg, neurological, structural, trauma).⁷⁻⁹

Do existing frameworks satisfy methodological guidelines for classification system development?

Overall, most studies identified by this review when rated using the Buchbinder risk of bias assessment tool¹² were considered low quality and the majority represented the opinion of a single author or author group, or they described a single case series with a limited description of how patients could be classified in their framework. Three papers were rated moderate quality using a specific risk-of-bias tool designed to assess the methodology of classification frameworks.^{7,9,18} Two of the sources described the use of consensus methodology in the development of their framework,^{7,9} and one of these studies provided additional inter-rater reliability scores to demonstrate validity of their classification groups.⁷

How conditions are classified

Different methodology was used by the authors for classifying voice disorders across the sources, as demonstrated in Table 3a, Table 3b, Table 3c, Table 4a and Table 4b. The multi-axis frameworks with fewest classification groups classified conditions according to broad aetiologies and their homogenous clinical features (Table 3a). The sub-categories and conditions included under each classification group appeared mutually exclusive, defined by clinical observations from multidimensional voice assessments.^{7,18} This was also a common theme in many of the single-axis frameworks (Table 4a). Frameworks where conditions were grouped by the presence or absence of organic pathology did not appear to include all possible conditions and the categories were not mutually exclusive (Table 3b).¹⁹⁻²¹ For example, conditions defined by psychological disturbances or those caused by vocal strain were included in categories labelled *Organic and Primary Organic*. Two authors chose to classify conditions according to the characteristics of underlying aetiology and disease (Table 3c). These contained the greatest number of classification groups and conditions were placed together in line with the location and characteristics of the organic pathology. This may be useful when identifying and communicating the specific organic cause of a presenting voice disorder, where an organic cause exists. However, in these frameworks the categories for voice disorders with no organic cause were not as clearly defined and the labelling of these categories within these frameworks was not as clear, for example *other disorders affecting voice*.⁹

Which conditions are included and what clinical assessments determine classification?

A wide range of disorders were described under each of the classification groups across the multiple frameworks in our review. The heterogeneity in terminology used for these conditions across the sources made comparison very difficult.

We have attempted to identify common themes for how these conditions were grouped. The most common themes included sub-classifying conditions with specific organic aetiologies, and sub-classification of observed vocal fold mucosal changes where muscle tension is considered the primary aetiology.

Due to a lack of detailed information contained within the sources described in this review, we could not address the sub-question of which clinical assessments are described to guide classification. However, the most frequently described clinical assessments included laryngeal examination through indirect laryngoscopy, auditory perceptual features of the voice and patient reported case history information.

Purpose of a classification framework for voice disorders

When defining a classification framework, it is important to consider the differences between classification and diagnosis, as well as their intended purpose. *Diagnostic criteria* are the set of signs, symptoms and tests, clinicians use to guide the process of diagnosis. These should have high levels of sensitivity to identify the different characteristics of a specific disease, whilst accurately distinguishing these features from other diseases.^{35,36} *Classification criteria* on the other hand, are commonly a set of standardized definitions, including signs symptoms and assessment measures, which describe broader well-defined homogenous cohorts of diseases, whilst identifying shared features of a condition.³⁷ The key difference is classification criteria do not include the full spectrum of indications of a disease, but should still have a set of criteria to identify a high proportion of patients and exclude a high proportion with other diseases.³⁶ An internationally approved classification framework should therefore be well accepted, unambiguous, encompass all domains of voice disorders, and be developed using appropriate research designs including consensus methodology, or inter-rater reliability measures.¹²

A suggested classification system

Having reviewed the sources of evidence selected against the review questions, there is no single classification framework that provides well defined labels and definitions to classify all the domains of voice disorders across the spectrum. We have summarised the evidence and common themes from the literature, and have proposed the following terms and defining criteria taken from a selection of existing frameworks be used for the purposes of voice disorder classification in clinical practice and research (Supplementary Figure, Infographic).

Muscle tension voice disorder/muscle tension dysphonia

The current review considers *Muscle Tension Dysphonia* (MTD) or *Muscle tension voice disorder* the most frequently used and well-defined term for voice disorders with hyper-functional-muscle tension aetiologies. Additionally, the

term is frequently cited in the current voice disorder literature demonstrating it has widespread acceptability. Furthermore, this review supports including sub classifications of benign pathology of the vocal folds when they are clearly identified as contributory or resulting from vocal hyperfunction.²⁵ This includes the prefix primary to denote no observable vocal fold changes, secondary to define a MTD with visible vocal fold trauma, and adaptive to define MTD in compensation to a primary organic aetiology.^{7,9} Classification criteria described in the published frameworks exclusive to this group of voice disorders include presenting signs or symptoms consistent with muscle hyperfunction in response to vocal demands of the individual^{7,9,27}; observed persistent laryngeal and/or supraglottic muscle constriction during phonation^{7,9,18}; improved voice, laryngeal observations, or perceived effort with specific targeted postures or reflex activities.^{7,18}

Functional (psychogenic) voice disorder

The two most frequently used terms for voice disorders of psychosocial aetiology, are *psychogenic* and *functional*. Our proposal is to adopt the term *functional* and the classification criteria as it pertains to voice disorders, as proposed in the most recent body of literature on Functional Neurological Disorders (FND).³⁸ The term *functional* has wide acceptance by neurologists and psychiatrists to classify disorders with motor and/or sensory signs or symptoms not explained by recognised neurological disorders. Recently an international multidisciplinary group of experts have agreed FND is a widely accepted classification term for a range of functional conditions, including dysphonia.³⁸ In addition, we propose that keeping the term *psychogenic* may help to avoid confusion, where historically *functional* has been used by SLPs and ENTs as a classification term encompassing both muscle tension and psychogenic aetiologies.

This review suggests we should avoid classification of these conditions using a process of diagnostic exclusion, as these voice disorders may not always be defined exclusively through a history of known conflicts or stressors in the presence of specific vocal characteristics and visible laryngeal tension.^{7,30} However, several exclusive features defined by Baker, Butcher and Aronson can be used to clearly identify them from other groups, most notably loss of voluntary motor control on the vocal sub-systems or an individual's inability to self-regulate initiation of optimal voice, albeit involuntary normal function may frequently be observed.^{7,18,30}

In our review, the voice disorder commonly termed *puberphonia* or *adolescent transitional dysphonia* was the only condition that could not be placed into our provisional analysis domains. We have grouped this condition within the *functional (psychogenic)* classification group based on classification criteria provided by Baker and the ASHA SIG-3 group.^{7,9} However, we acknowledge the aetiology of this disorder can be multifactorial and may not always meet the criteria of a functional disorder. For the purposes of this review, we have placed it

as a sub-group with criteria that identifies this condition by non-organic aetiology, that is, in the presence of normal hormonal development.

Organic voice disorders defined by etiological group

Diagnostic criteria for specific voice disorders or diseases of the vocal tract that arise from organic or neurogenic causes are frequently well-defined in the literature. However, there remains a need for commonly understood classification terminology where groups of homogeneous disorders with similar defining criteria can be placed. Existing frameworks most frequently group conditions according to broad aetiologies and the area of breakdown within the upper aerodigestive tract. The most comprehensive example is detailed in the classification manual for voice disorders-I (CMVD-I).⁹ This model demonstrated a high-quality rating in our assessment, is easy to follow, and the disorders are grouped appropriately to their underlying aetiology where organic conditions are concerned. For each classification group the model provides a comprehensive list of specific conditions with well-defined organic or neurogenic pathophysiology. In addition, the frameworks for BVFL and vocal fold immobility^{32,33} are examples of classification systems where specific terminology and clearly defined criteria can be used to aid the communication of both diagnosis and treatment planning.

Pathway to a universal classification system

Based on this review, more research is needed to establish a universally accepted classification system in voice disorders. The necessary steps to this outcome would include an international multi-disciplinary approach using well established consensus methodology such as a Delphi process. The aim would be to reach agreement on i) the terminology to label common groups of voice disorders for the purpose of classification; ii) nomenclature for specific diagnoses or conditions of voice disorders that are often poorly defined in the literature (eg, FND, Puberphonia); iii) high quality research using standardised assessment approaches to define the common clinical features of all voice disorders, with measures of inter-rater reliability to enable consistent classification under an agreed framework. A framework with international acceptance could promote international collaboration and greatly improve evidence-based clinical research in the field of voice disorder diagnostics, as well as improving communication between professionals when planning treatment for voice disorders.

Limitations

To our knowledge, this is the first scoping review to examine the frameworks and terminology used to define voice disorders across the full range of aetiologies within the health sciences literature. Nevertheless, there are some limitations to our review. We only reviewed English language studies, however very few non-English papers were identified in our searches, and thus unlikely

to have had a substantial impact on the findings of our review. We incorporated critical appraisal and risk of bias on selected multi-axis frameworks to facilitate a quality assessment, but acknowledge this is not a requirement or in some examples a recommendation of scoping reviews. Unlike systematic reviews that are aimed at a narrow range of evidence and may require quality assurance, scoping reviews are less concerned with a specific question and the sources of evidence are from a much broader search.³⁹ The decision to include a quality assessment was considered important to help determine a core set of classification terminology and criteria, which was an objective of our review.

CONCLUSIONS

Many classification frameworks have been proposed for use with disorders of voice and have been summarised as part of this scoping review in the endeavour to propose a framework for clinical and research use in the future. Most frameworks are of low-quality evidence with no discussion of the method of how they have been devised. Almost half of the terms presented have poorly defined classification criteria and are not widely used within the current scientific literature. No single classification system provides clearly defined distinctions across the full spectrum of aetiologies. However, several key themes are presented together with a core set of suggested terms and definitions for the broad classification of voice disorders. We propose a classification framework which will require future international and interdisciplinary consultation, and evaluation. Continued high quality research studies using consensus methodology and inter-rater reliability scores is recommended to determine agreement of the labels used to classify homogenous groups of voice disorders and determine inclusion of specific conditions under each of the classification groups.

DECLARATIONS OF COMPETING INTEREST

None

Acknowledgments

The authors would like to acknowledge the assistance and expertise of Ms. Elaine Tam, Librarian for the Faculty of Medicine and health, The University of Sydney, and Ms. Sarah Thorning, Librarian at Gold Coast University Hospital.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at [doi:10.1016/j.jvoice.2022.02.009](https://doi.org/10.1016/j.jvoice.2022.02.009).

APPENDIX I: SEARCH STRATEGY

MEDLINE (Ovid SP) Search conducted on February 24, 2020

Search	Query	Records retrieved
#1	Voice disorders/OR Voice disorder*.mp OR Laryngeal disorder*.mp OR Vocal disorder*.mp OR Vocal fold*.mp OR Dysphonia.mp OR Organic voice disorder*.mp OR Organic dysphonia*.mp OR larynx* structural.mp OR larynx* inflammation.mp OR laryngeal disease*.mp OR Functional voice disorder*.mp OR Functional dysphonia*.mp OR Functional-muscle tension voice disorder*.mp OR Muscle tension dysphonia*.mp OR Muscle tension voice disorder*.mp OR Muscle misuse dysphonia*.mp OR Muscle misuse voice disorder*.mp OR Ventricular dysphonia*.mp OR Non organic voice disorder*.mp OR Functional-psychogenic voice disorder*.mp OR Psych* voice disorder.mp OR Psych* voice disorders.mp OR Functional neurological voice disorder*.mp OR Neuro* voice disorder.mp OR Neuro* voice disorders.mp OR Neuro* dysphonia.mp or Neuro* dysphonia.mp	21235
#2	Description.mp OR Descriptor*.mp OR "international classification of diseases"/ OR Classification*.mp OR Classification/ OR Classification criteria.mp OR Classification terminology.mp OR Framework*.mp OR Practice guideline/OR Guideline*.mp	1684932
#3	1 AND 2	1400
	Limited to (english language and humans and yr = "1900-current")	1041

REFERENCES

- Zhang Z. Mechanics of human voice production and control. *J Acoust Soc Am*. 2016;140:2614–2635. <https://doi.org/10.1121/1.4964509>.
- Merrill RM, Roy N, Lowe J. Voice-related symptoms and their effects on quality of life. *Ann Otol Rhinol Laryngol*. 2013;122(6):404–411. <https://doi.org/10.1177/000348941312200610>.
- Benninger MS, Holy CE, Bryson PC, et al. Prevalence and occupation of patients presenting with dysphonia in the United States. *J Voice*. 2017;31(5):594–600. <https://doi.org/10.1016/j.jvoice.2017.01.011>.
- Stachler RJ, Francis DO, Schwartz SR, et al. Clinical practice guideline: hoarseness (Dysphonia) (Update). *Otolaryngol Head Neck Surg*. 2018;158(1_suppl):S1–S42. <https://doi.org/10.1177/0194599817751030>.
- Roy N. Functional dysphonia. *Curr Opin Otolaryngol Head Neck Surg*. 2003;11(3):144–148.
- Cohen SM, Jaewhan K, Roy N, et al. Delayed otolaryngology referral for voice disorders increases health care costs. *Am J Med*. 2015;128(4):426.e11–426.e18. <https://doi.org/10.1016/j.amjmed.2014.10.040>.
- Baker J, Ben-Tovim DI, Butcher A, et al. Development of a modified diagnostic classification system for voice disorders with inter-rater reliability study. *Logoped Phoniatr Vocol*. 2007;32(3):99–112. <https://doi.org/10.1080/14015430701431192>.
- Bradley PJ, et al. Voice disorders: classification. In: Anniko M, ed. *Otorhinolaryngology, Head and Neck Surgery*. Berlin: Springer-Verlag Berlin; 2010:555–562.
- Verdolini K, Rosen C, Branski R, et al. *Classification Manual for Voice Disorders-I (1st ed.)*. New York: Psychology Press; 2005.

10. Roy N, Barkmeier-Kraemer J, Eadie T, et al. Evidence-based clinical voice assessment: a systematic review. *Am J Speech Lang Pathol.* 2013;22(2):212–226. [https://doi.org/10.1044/1058-0360\(2012\)12-0014](https://doi.org/10.1044/1058-0360(2012)12-0014).
11. Yazici H. Diagnostic versus classification criteria: a continuum. *Bull NYU Hosp Jt Dis.* 2009;67(2):206–208.
12. Buchbinder R, Goel V, Bombardier C, et al. Classification systems of soft tissue disorders of the neck and upper limb: do they satisfy methodological guidelines? *J Clin Epidemiol.* 1996;49(2):141–149.
13. Tricco AC, Lillies E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med.* 2018;169:467–473. <https://doi.org/10.7326/M18-0850>.
14. Peters MDJ GC, McInerney P, Baldini Soares C, et al. Chapter 11: scoping reviews. In: Aromataris E, Munn Z, eds. *Joanna Briggs Institute Reviewer's Manual*. The Joanna Briggs Institute; 2017. Available at; <https://reviewersmanual.joannabriggs.org/>.
15. Payten CL, Chiapello G, Weir KA, et al. Terminology and frameworks used for the classification of voice disorders: a scoping review protocol. *JBI Evidence Synthesis.* 2020;19:454–462. <https://doi.org/10.11124/JBIES-20-0006616>.
16. Bramer WM, Giustini D, de Jonge GB, et al. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc.* 2016;104(3):240–243. <https://doi.org/10.3163/1536-5050.104.3.014>.
17. Covidence Systematic Review Software, V.H.I., Melbourne, Australia. Available at: www.covidence.org.
18. Morrison MD, Nichol H, Rammage LA. Diagnostic criteria in functional dysphonia. *Laryngoscope.* 1986;96(1):1–8. <https://doi.org/10.1288/00005537-198601000-00001>.
19. Rosen CA, Murry T. Nomenclature of voice disorders and vocal pathology. *Otolaryngol Clin North Am.* 2000;33(5):1035–1046.
20. Damste PH. Vocal disorders. A guide to their diagnosis. *ORL J Otorhinolaryngol Relat Spec.* 1973;35:149–153.
21. Milutinovic Z. Classification of voice pathology. *Folia Phoniater Logop.* 1996;48:301–308.
22. Aronson AE, Peterson Jr. HW, Litin EM. Psychiatric symptomatology in functional dysphonia and aphonia. *J Speech Hear Disord.* 1966;31:115–127.
23. Koufman JA, Blalock PD. Classification and approach to patients with functional voice disorders. *Ann Otol Rhinol Laryngol.* 1982;91(4 Pt 1):372–377.
24. Hillman RE, Holmberg EB, Perkell JS, et al. Objective assessment of vocal hyperfunction: an experimental framework and initial results. *J Speech Hear Res.* 1989;32:373–392.
25. Hillman RE, Stepp CE, Van Stan JH, et al. An updated theoretical framework for vocal hyperfunction. *Am J Speech Lang Pathol.* 2020;29(4):2254–2260. https://doi.org/10.1044/2020_AJSLP-20-00104.
26. Morrison MD, Rammage LA. Muscle misuse voice disorders: description and classification. *Acta Otolaryngol.* 1993;113:428–434.
27. Van Houtte E, Van Lierde K, Claeys S. Pathophysiology and treatment of muscle tension dysphonia: a review of the current knowledge. *J Voice.* 2011;25(2):202–207. <https://doi.org/10.1016/j.jvoice.2009.10.009>.
28. Spencer ML. Muscle tension dysphonia: a rationale for symptomatic subtypes, expedited treatment, and increased therapy compliance. *Perspect Voice Voice Disord.* 2015;25(1):5–15. <https://doi.org/10.1044/vvd25.1.5>.
29. Fernández S, Garaycochea O, Martínez-Arellano A, et al. Does more compression mean more pressure? A new classification for muscle tension dysphonia. *J Speech Lang Hear Res.* 2020;63(7):2177–2184. https://doi.org/10.1044/2020_JSLHR-20-00042.
30. Butcher P, Elias A, Cavalli L. *Understanding and Treating Psychogenic Voice Disorder: A CBT Framework*. Wiley Series in Human Communication Science. New York, NY: John Wiley & Sons Ltd; 2007.
31. Hacki T, Moerman M, Rubin JS. Malregulative' rather than 'Functional' dysphonia: a new etiological terminology framework for phonation disorders—a position paper by the Union of European Phoniaticians (UEP). *J Voice.* 2022;36(1):50–53. <https://doi.org/10.1016/j.jvoice.2020.04.032>.
32. Rosen CA, Mau T, Hess M, et al. Nomenclature proposal to describe vocal fold motion impairment. *Eur Arch Otorhinolaryngol.* 2016;273:1995–1999. <https://doi.org/10.1007/s00405-3663-0>.
33. Rosen CA, Gartner-Schmidt J, Hathaway B, et al. A nomenclature paradigm for benign midmembranous vocal fold lesions. *Laryngoscope.* 2012;122:1335–1341.
34. Arnold GE, Pinto S. Ventricular dysphonia: new interpretation of an old observation. *Laryngoscope.* 1960;70(12):1608–1627. <https://doi.org/10.1288/00005537-196012000-00003>.
35. Aggarwal R, Ringold S, Khanna D, et al. Distinctions between diagnostic and classification criteria? *Arthritis Care Res.* 2015;67(7):891–897. <https://doi.org/10.1002/acr.22583>.
36. Hunder GG, Arend WP, Bloch DA, et al. The American College of Rheumatology 1990 criteria for the classification of vasculitis: introduction. *Arthritis Rheum.* 1990;33(8):1065–1067.
37. Hunder GG. The use and misuse of classification and diagnostic criteria for complex diseases. *Ann Intern Med.* 1998;129(5):417–418. <https://doi.org/10.7326/0003-4819-129-5-199809010-00013>.
38. Baker J, Barnett C, Cavalli L, et al. Management of functional communication, swallowing, cough and related disorders: consensus recommendations for speech and language therapy. *J Neurol Neurosurg Psychiatr.* 2021;92(10):1112–1125. <https://doi.org/10.1136/jnnp-2021-326767>.
39. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005;8(1):19–32. <https://doi.org/10.1080/1364557032000119616>.