Review

A Scoping Literature Review of Relative Fundamental Frequency (RFF) in Individuals with and without Voice Disorders

Victoria S. McKenna [1,2,3,*], Jennifer M. Vojtech [4,5], Melissa Previtera [6], Courtney L. Kendall [1] and Kelly E. Carraro [1]

1   Department of Communication Sciences and Disorders, University of Cincinnati, Cincinnati, OH 45267, USA
2   Department of Biomedical Engineering, University of Cincinnati, Cincinnati, OH 45219, USA
3   Department of Otolaryngology–Head & Neck Surgery, University of Cincinnati, Cincinnati, OH 45267, USA
4   Delsys, Natick, MA 01760, USA
5   Altec, Natick, MA 01760, USA
6   Health Sciences Library, University of Cincinnati, Cincinnati, OH 45267, USA
*   Correspondence: mckennvs@ucmail.uc.edu

Abstract: Relative fundamental frequency (RFF) is an acoustic measure that characterizes changes in voice fundamental frequency during voicing transitions. Despite showing promise as an indicator of vocal disorder and laryngeal muscle tension, the clinical adoption of RFF remains challenging, partly due to a lack of research integration. As such, this review sought to provide summative information and highlight next steps for the clinical implementation of RFF. A systematic literature search was completed across 5 databases, yielding 37 articles that met inclusion criteria. Studies most often included adults with and without tension-based voice disorders (e.g., muscle tension dysphonia), though patient and control groups were directly compared in only 32% of studies. Only 11% of studies tracked therapeutic progress, making it difficult to understand how RFF can be used as a clinical outcome. Specifically, there is evidence to support within-person RFF tracking as a clinical outcome, but more research is needed to understand how RFF correlates to auditory-perceptual ratings (strain, effort, and overall severity of dysphonia) both before and after therapeutic interventions. Finally, a marked increase in the use of automated estimation methods was noted since 2016, yet there remains a critical need for a universally available algorithm to support widespread clinical adoption.

Keywords: acoustics; voice; review

1. Introduction

Relative fundamental frequency (RFF) is an acoustic measure that has gained substantial traction for its use in clinical voice evaluation. Over the last two decades, research into RFF has yielded evidence to suggest that this non-invasive, objective measure may assist in assessing and tracking the degree of laryngeal muscle tension [1] and its associated symptoms [2]. It has been estimated that 65% of patients with voice disorders have excessive laryngeal tension [3], making RFF an appealing acoustic measure for clinical diagnosis and tracking.

Within the literature, RFF has been shown to distinguish between typical voices and those characterized by excessive laryngeal muscle tension (i.e., tension-based voice disorders), such as laryngeal dystonia [4], Parkinson's disease (PD) [5], and vocal hyperfunction (VH) [6]. Further, RFF is significantly correlated with vocal effort [4,7,8], as well as perceived dysphonia and/or strain [2,9].

Despite the promising results demonstrated by these studies, RFF is not yet a staple of standardized voice assessments. The discrepancy between research and clinical support for RFF may be, in part, due to a lack of clear understanding of its application and interpretation across patient populations. Yet, a comprehensive review of the RFF literature has not been undertaken and, moreover, summative information for RFF as a clinical tool is

nonexistent. Thus, the purpose of this scoping review was to carry out an exhaustive search into the existing RFF literature to be able to condense current knowledge and provide recommendations for next steps needed toward the clinical adoption and interpretation of RFF.

1.1. Stimuli and Calculation

Traditional acoustic measures of voice—such as jitter and shimmer—are computed during voicing steady-states, most often during sustained vowel productions. RFF is a unique voice measure because it quantifies changes in fundamental frequency (f_o) during voicing *transitions*. Specifically, RFF is measured from the instantaneous changes in f_o during vocalic devoicing and voicing gestures during a vowel–voiceless consonant–vowel (VCV) utterance. Usual RFF stimuli include VCV utterances of /ifi/, /ɑfɑ/, and /ufu/ [10], or specific wording from the phonetically balanced Rainbow Passage [11]. For example, the VCV transition within the word "beautiful" (i.e., /ɪfə/) meets the criteria for RFF extraction.

RFF is calculated from the 10 voicing cycles immediately before the voiceless consonant ("offset cycles") and 10 voicing cycles immediately after the voiceless consonant ("onset cycles"). As such, an estimate of RFF for one speaker comprises 20 individual values; each value corresponds to a change in f_o from steady state and can be analyzed either alone as an estimate of laryngeal tension or alongside other values to understand trends in the speaker's devoicing and/or voicing gestures. When reporting RFF findings, individual voicing cycle values are labeled sequentially from 1 to 10 as they occur in time, with offset cycle 10 and onset cycle 1 located closest to the consonant (see Figure 1 for example of labeled offset and onset cycles).
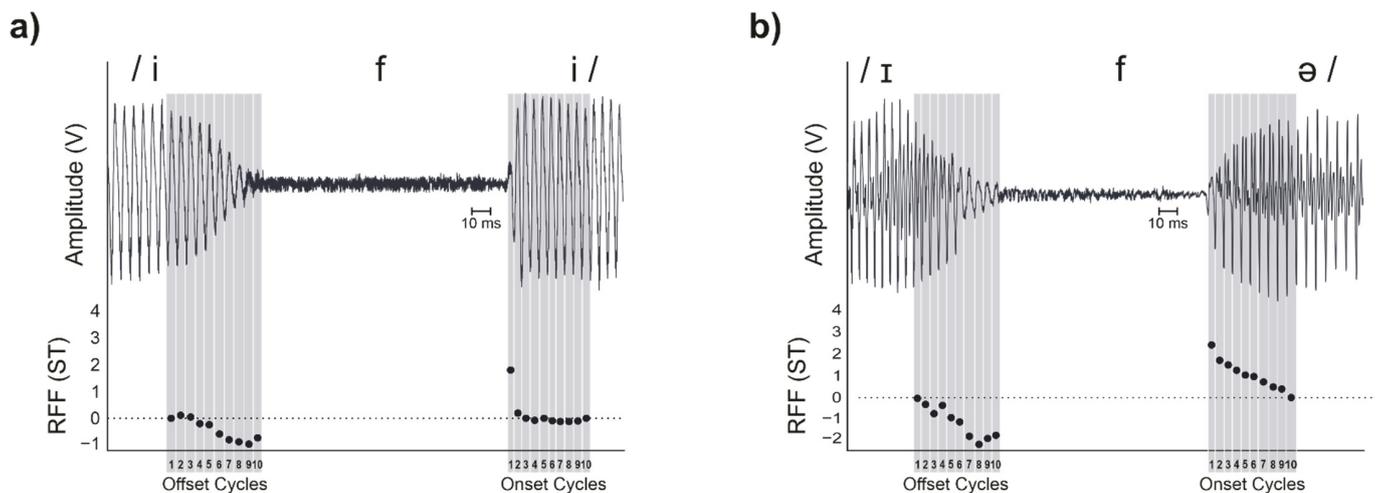
Figure 1. Offset and onset cycle identification and RFF cycle values for (**a**) the VCV utterance /ifi/, and (**b**) the portion /ɪfə/ in the word "beautiful".

RFF values for a given voicing cycle are determined by first calculating the instantaneous f_o of the cycle (i.e., the inverse of the cycle period) in units of Hertz (Hz). The cycle f_o is normalized to that of the voicing cycle closest to the steady-state portion of its corresponding vowel ($f_o{}^{ref}$); for voicing offset cycles, this is the first voicing cycle (cycle 1) in the first vowel of the VCV utterance and for voicing onset cycles, this is the last voicing cycle (cycle 10) of the second vowel. Finally, all cycle values (in Hz) are converted to semitones (ST), as shown in Equation (1).

$$\text{RFF (ST)} = 12 \times \log_2\left(\frac{f_o}{f_o{}^{ref}}\right) \tag{1}$$

Normalizing RFF values from Hz to ST helps to standardize the measurement to account for a speaker's own variations in f_o production across vowels, and additionally allows for comparisons across speakers. Therefore, RFF cycle values reflect a standardized change in the frequency of cycle vibration from the steady state of each vowel. Specifically, an RFF cycle value of 0 ST reflects no change in vibratory timing from the vowel steady state, whereas a positive RFF cycle value indicates higher f_o value relative to steady state (i.e., faster vocal fold vibration) and a negative RFF cycle value indicates lower f_o value relative to steady state (i.e., slower vocal fold vibration).

Importantly, Watson [12] was the first to publish on the RFF measure using the normalization and conversion procedures explained above. At the time, he referred to the measures only as "cycle values during devoicing." It was not until Stepp et al. [13] that the term "relative fundamental frequency" was coined, and subsequently adopted across the voice literature.

1.2. Physiological Basis

Prior to the development of the RFF metric, several researchers examined cycle-specific changes in f_o to describe phonatory offset and onset behaviors in similar phonetically constrained contexts. It has been postulated that two primary laryngeal adjustments occur to cease vocal fold vibration during intervocalic offsets: vocal fold abduction and increased laryngeal muscle tension [14]. Vocal fold abduction is primarily attributed to the posterior cricoarytenoid muscle and has been shown to reduce transglottal pressure as well as reduce the duration of vocal fold contact [15] to inhibit vocal fold vibration. At the same time, increased laryngeal muscle tension from increased activity of the cricothyroid (CT) [16] and vocalis [17] muscles are suspected to help cease vibration, though there is additional evidence to suggest that overall laryngeal height and/or tilt (mediated primarily by the extrinsic laryngeal musculature) may also play a role [18].

Intervocalic onset behavior is influenced by the preceding voiceless consonant [19,20], with laryngeal tension established during the voiceless consonant carrying over into the initial voicing cycles. The existing tension in the CT and vocalis are thought to increase f_o [16,21] while vocal fold adduction acts to raise transglottal pressure and assist in the re-initiation of voicing. Prior work has shown that initial voicing onset cycles exhibit relatively higher f_o when compared to subsequent voicing cycles [22,23].

Stepp et al. [24] integrated information from these previous works to propose a theoretical model to interpret RFF in speakers with and without voice disorders. The authors hypothesized that offset and onset RFF values rely on three laryngeal factors: kinematics, aerodynamics, and muscle tension. Intervocalic offset cycles generally stabilize around 0 ST in speakers with typical voices, indicating a balance between the increasing laryngeal muscle tension needed to stop vocal fold vibration (increasing vocal fold vibration and cycle speed) and reduced transglottal pressure due to vocal fold abduction for the voiceless consonant production (decreasing vocal fold vibration and cycle speed). Conversely, intervocalic onset cycles in speakers with typical voices are consistently more positive than offset values at the start of the second vowel, hovering between 2 and 3 ST for onset cycle 1. The high RFF values observed during initial onset cycles are thought to occur due to increased transglottal pressure and peak flow during adduction [20] in combination with elevated longitudinal tension carried over from the consonant production.

It follows that someone with a voice disorder who has impacted airflow, laryngeal kinematics, and/or laryngeal muscle tension would show aberrant RFF. Stepp et al. [24] hypothesized that patients with voice disorders characterized by excessive tension of the intrinsic and/or extrinsic laryngeal musculature would exhibit lower RFF values. Specifically, the authors proposed that elevated laryngeal muscle tension at baseline would prohibit increases in tension levels during voicing offset, which would typically help to counteract the effects of abduction. A lack of increased tension from baseline was speculated to result in more negative RFF offset cycle values compared to adults with typical voices, where values hover around 0 ST for voicing offset. It was also hypothesized that voicing

onset would be affected by this inability to further modulate laryngeal tension, resulting in lower RFF onset values when compared to adults with typical voices.

Heller Murray et al. [6] expanded on this model by distinguishing the effects of longitudinal and transverse laryngeal muscle tension for specific patient populations. The authors proposed that patients with phonotraumatic lesions may have elevated longitudinal and transverse tension—or tight adduction during phonation [25]—that interacts with the laryngeal kinematics proposed in the Stepp et al. [24] model. It was hypothesized that additional transverse vocal fold tension would increase vocal fold contact time, thereby reducing the effects of vocal fold abduction and consequently lowering RFF values beyond the effects of longitudinal tension alone. Likewise, increased transverse tension was thought to reduce the duration of the adductory gesture and inhibit the impact of aerodynamic forces on RFF to, once again, lower RFF values as compared to patients with only elevated longitudinal tension at baseline. This model was supported via evidence of significantly lower RFF offset and onset cycle values for the voicing cycles located closest to the voiceless consonant (i.e., offset cycle 10 and onset cycle 1) in patients with phonotraumatic lesions when compared to those with non-phonotraumatic VH.

1.3. Clinical Challenges

Despite growing theoretical and empirical evidence for RFF as an acoustic indicator of laryngeal muscle tension, RFF is not included within the battery of acoustic metrics recommended for standardized clinical voice evaluations [26]. Two of the primary barriers inhibiting the clinical implementation of RFF include the extensive training required to reliably identify RFF cycles during voicing transitions, as well as the time-consuming nature of manually extracting RFF values from the acoustic signal. It has been shown that at minimum, six RFF extractions are needed to establish a consistent and reliable estimate for a single speaker [4]. With the time demands placed on clinical staff, an extraction time of 20–30 min for a single acoustic measure is not feasible. The implementation of an algorithm to extract estimates of RFF is an attractive alternative to the time- and training-intensive nature of manual RFF estimation and, moreover, is vital for clinical adoption.

In addition to faster processing methods, more information is needed on how to interpret RFF values in the context of disease-specific processes and treatment outcomes. RFF offset cycle 10 and onset cycle 1 have received the most attention in the RFF literature as potential clinical markers. It has been hypothesized that RFF values at these cycles are the most sensitive to the aerodynamic, tension, and kinematic factors proposed in the theoretical models due to their location furthest from the steady-state of each vowel (and closest to the voiceless consonant). Although studies have reported significant differences between individuals with and without tension-based voice disorders [5,6,27], there is still no universally accepted clinical cut-off value for offset cycle 10 or onset cycle 1 to distinguish between vocal health and disorder.

1.4. Purpose

To our knowledge, no study has formally undertaken a review of the RFF literature to integrate findings and synthesize recommendations for clinical adoption. Therefore, we completed a scoping review to identify and describe the current published literature on RFF. Through this review, we aimed to: (i) summarize the methodology for acquiring and processing RFF measures, (ii) compile reported RFF offset 10 and onset 1 values, (iii) describe statistical comparisons and outline relationships with perceptual measures, and (iv) provide recommendations for future steps towards clinical implementation and interpretation.

2. Materials and Methods

A scoping review is a systematic way to summarize findings and identify gaps in literature across studies with heterogenous methods. A scoping review was chosen because it seemed most appropriate for undertaking a comprehensive review of all RFF literature, without limitation to specific patient populations, study designs, or method-

ological approaches. We followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) checklist as outlined in Tricco et al. [28], which identifies 20 essential reporting items.

2.1. Literature Search

A systematic search of the literature was facilitated by a library scientist (author M.P.) on February 26, 2020. Five databases were searched: Embase (Elsevier, Amsterdam, The Netherlands), Medline (EBSCOhost), PsychInfo, CINAHL, and INSPEC. No filters or time limits were applied to the search. An example search string can be found in Appendix A, but generally included the following inputs for each database, modified as appropriate: relative fundamental frequency, fundamental frequency and (i) devoicing, (ii) offset, (iii) onset, (iv) voiceless obstruent, and (v) cycle-to-cycle. At the same time, a grey literature search was conducted by contacting a lead researcher in the field to share manuscripts that had been peer-reviewed and accepted for publication, but not yet formally published. All identified articles were first saved to EndNote where de-duplication began. Files were transferred to a review software management system, known as Rayyan [29], wherein any remaining duplicates could be identified and removed.

To ensure that this scoping review was as current as possible, a second search was completed on 9 April 2021. The same databases and search strings were used but the dates were limited to those between the first and second searches (i.e., 26 February 2020 to 9 April 2021). Finally, a hand search of the literature that included accepted manuscripts that had not yet been published was also completed in September of 2021 to find any final studies.

2.2. Study Review Procedures

2.2.1. Study Inclusion and Exclusion Criteria

Inclusion and exclusion criteria were set prior to study review. Inclusion criteria were as follows: (i) enrollment of human subjects, (ii) entire study reported in the English language, and (iii) analysis of relative fundamental frequency. Exclusion criteria were as follows: (i) conference abstracts, (ii) case study or single-case design, (iii) meta-analysis or review, or (iv) studies where voicing cycles were not normalized to ST. We did not limit our inclusion by subject age, medical status (typical voice, disordered voice), or the type of equipment used to acquire voice acoustics (e.g., microphone, neck-surface accelerometer).

2.2.2. Article Review and Data Extraction

Prior to the official title/abstract review, two authors (V.M., C.K.) independently reviewed a random selection of 200 titles/abstracts from the literature search. This reliability review yielded excellent reliability with 99.5% absolute agreement. Subsequently, the two authors then divided the title/abstract review, wherein each author completed a review of half of the identified studies.

Following the title/abstract screening, the two authors independently read each study in full, resulting in no discrepancies between independent decisions, or 100% agreement for final inclusion in this review. This same process was completed in April of 2021 during the secondary database review and in September of 2021 for the final hand search.

Once included studies were identified, authors independently extracted information and placed all data into a structured Excel (Microsoft; Redmond, WA, USA) document. Author V.M. extracted information from 100% of the included studies, whereas author C.K. extracted information from 50% of the studies and author K.C. extracted the remaining 50% of the studies. Researchers were blind to each other's extractions. Then, the researchers met and reviewed their extractions together and discussed any discrepancies to consensus. The information of interest for this review included the following:

1. Study Information: author(s), year published, journal
2. Subject Information: number of subjects, ages, sex/gender, diagnoses
3. RFF Extraction Methods: speech stimuli, processing methods

4. RFF Values for Offset Cycle 10 and Onset Cycle 1: for each population and/or study condition. Values that were reported in graphical form were not included or estimated.
5. Statistical Results and Interpretations.

3. Results

Our literature search (Search #1, February 2020), secondary search (Search #2, April 2021), and hand search (September 2021) yielded 10,806 studies for review. Once duplicates were removed, 5613 studies were reviewed by title and abstract only. From here, 51 studies were read in full. A total of 14 studies were excluded at this point due to having the wrong measure (*n* = 12), not enrolling human subjects (*n* = 1), and being a conference abstract (*n* = 1). Finally, 37 studies met criteria for inclusion in the scoping review [1,2,4–10,12,13,24,27,30–53]. Please see Figure 2 for a flowchart on the review process.
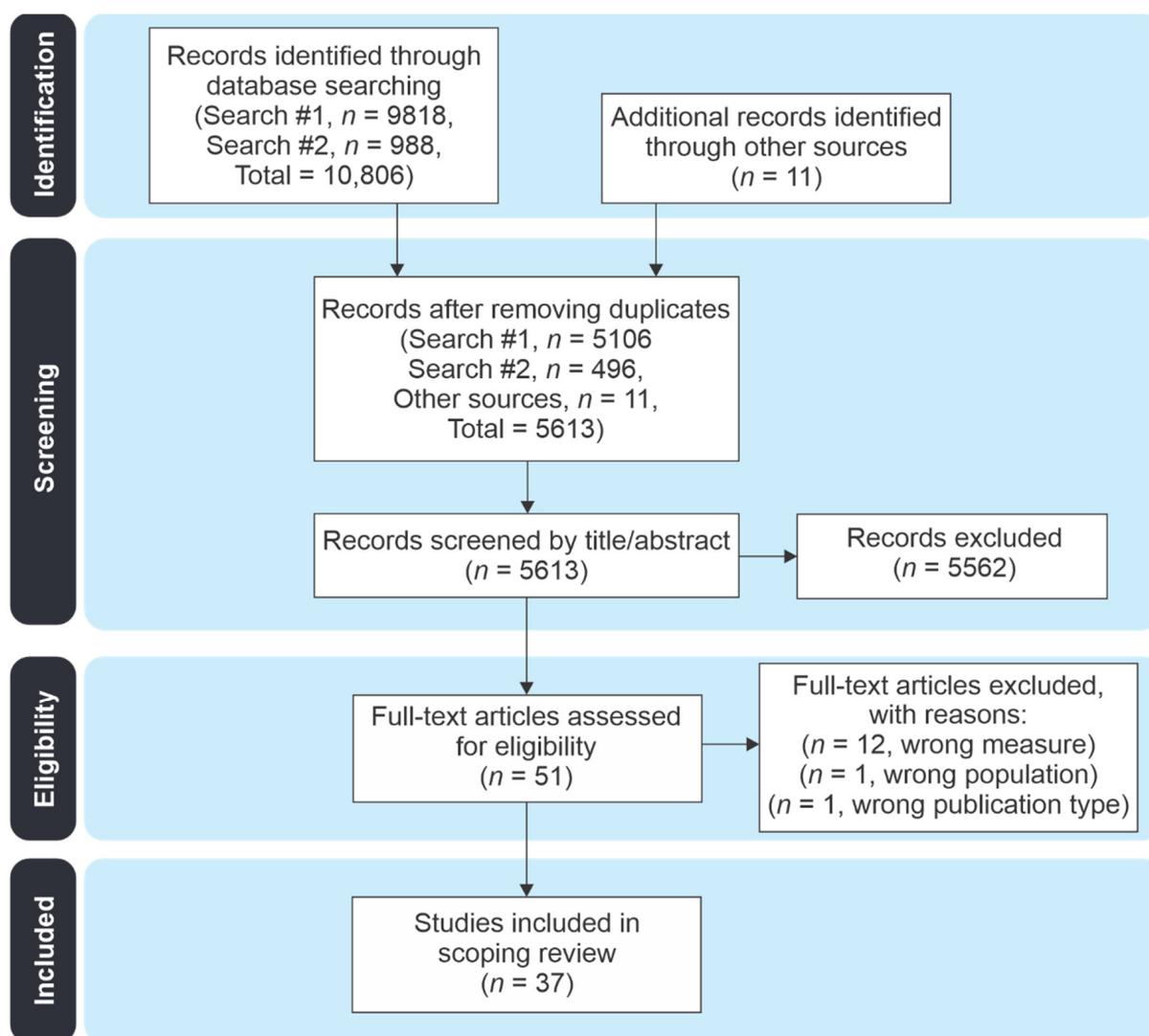
Figure 2. Flowchart of identified, screened, and selected studies.

3.1. Study Summary Information

Publication dates ranged from the years 1998 to 2022 (inclusive of in-press publications identified in 2021 and published in 2022) with most of the studies being published post 2009 (*n* = 34). The majority of studies enrolled vocally healthy adults with typical voices (*n* = 29; 78%), though both aging adults (*n* = 9; 24%) and children (*n* = 3; 8%) have been

examined. A total of 16 studies (43%) enrolled subjects with voice disorders, with the most common diagnoses being muscle tension dysphonia (MTD; $n = 11$; 30%), phonotraumatic lesions ($n = 11$; 30%), PD ($n = 8$; 22%), and laryngeal dystonia ($n = 6$; 16%). The number of subjects varied widely across studies from as low as 8 [9] to as high as 483 [51].

3.2. RFF Acquisition and Processing

The speech stimuli used to acquire RFF were evenly distributed across studies. A total of 19 studies (51%) analyzed VCV utterances and 16 studies (43%) analyzed RFF from continuous speech. Only two studies (5%) employed both VCV and continuous speech in their methodologies. The intervocalic consonant /f/ results in the lowest within-subject variation in RFF measures [38] and was unsurprisingly the most common voiceless consonant employed in RFF VCV stimuli.

Five studies focused on algorithmic development for automated RFF extraction [33,39,40,51,52]. Semi-automated algorithmic development began in 2013, in which a custom MATLAB (The MathWorks, Natick, MA, USA) program extracted RFF values from VCV utterances recorded with a microphone [40]. The algorithm (aRFF) [39] was updated in 2017 and openly shared on Boston University's website at no cost to the user. Vojtech and colleagues [51] refined the algorithm via incorporating rule-specific criteria for tuning algorithmic parameters based on patient-specific voice characteristics and acoustic recording environment (referred to as aRFF-AP; current version) and, later, by cross-referencing simultaneous high-speed laryngeal imaging to hone in on cycle-specific extraction decisions around the voiceless consonant (referred to as aRFF-APH; unreleased) [52].

These semi-automated algorithms operate in six steps: (1) vowel and voiceless fricative detection via high-to-low energy ratios, (2) f_o estimation via autocorrelation (aRFF) or Auditory-SWIPE′ (aRFF-AP, aRFF-APH), (3) vocal cycle peak detection of the vowels, (4) vowel-fricative boundary detection via acoustic features, (5) sample rejection for instances that do not meet specified criteria (e.g., less than 10 onset or offset cycles, glottalization, misarticulation, voicing during the voiceless consonant), and (6) RFF calculation. Each step of the algorithm is fully automated except for step (1), in which the user must confirm or alter the algorithmically identified location of the fricative within the acoustic signal.

Of the five studies that focused on algorithmic development, four studies developed automated extraction methods for acoustic signals acquired with a microphone, whereas Groll et al. [33] developed automated methods from neck-surface accelerometer signals. Accelerometers are wearable sensors that have the potential to expand the use of RFF to ambulatory monitoring where it could have utility for long-term clinical tracking. Unlike the microphone-based algorithms, the accelerometer-based algorithm is fully automated and does not require user input. See Table S1 for more information on the studies devoted to automated RFF extraction.

A total of 32 studies in this review did not focus on automated RFF development. Of these, 23 studies (72%) used manual estimation to calculate RFF and 9 studies (28%), beginning in 2016, used semi-automated estimation methods. With manual estimation, voicing onset and offset cycles were identified through visual inspection of the acoustic waveform. Praat, a freely available acoustic software [54], was used in 74% ($n = 17$) of studies that used manual estimation. Praat's "pulse" function uses an autocorrelation method to determine cycle f_o from the acoustic waveform. Of the nine studies that used semi-automated estimation, eight employed the aRFF algorithm and only one used aRFF-AP. The aRFF algorithm also uses the autocorrelation method to determine cycle f_o values from the acoustic waveform. Given that manual RFF estimation via Praat and semi-automated RFF estimation via aRFF have been the most widely implemented methods, autocorrelation was the primary technique used to determine f_o in most published studies. See Table S2 for characteristics and data extraction from all 32 studies.

3.3. Onset Cycle 1 and Offset Cycle 10

A total of 14 studies reported offset cycle 10 values and 13 studies reported onset cycle 1 values for adults with typical voices under the age of 65 years. The distribution for offset cycle 10 ranged from −1.10 to 0.05 ST with a median value of −0.6 ST, whereas onset cycle 1 values ranged from 1.68 to 3.82 ST with a median value of 2.60 ST. See Figure 3 for a boxplot distribution of RFF values for adults with typical voices.
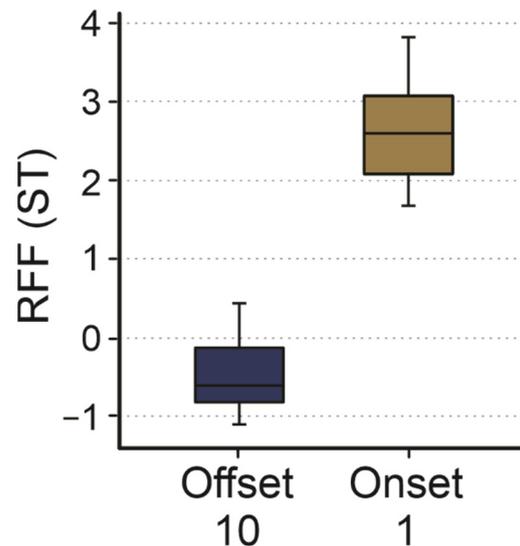

Figure 3. Boxplot of average RFF offset cycle 10 and onset cycle 1 values in studies on adults aged 18–65 years with typical voices.

Mean offset cycle 10 values were consistently lower for individuals with voice disorders when compared to those of typical voices. Individuals with adductor laryngeal dystonia had reported values of −1.20 ST [2,4], and those with PD exhibited −1.90 ST [49] and −2.20 ST [5]. Of the studies that reported values in patients with VH—including both phonotraumatic lesions and non-phonotraumatic VH subgroups—offset cycle 10 ranged from −1.76 ST to −0.80 ST, with a median value of −1.35 ST [6,13,24,27]. Two studies completed formalized receiver operating characteristic (ROC) curves to determine the RFF value that would most accurately distinguish between typical voices and those with VH. Stepp et al. [27] identified a value of −0.56 ST (97% sensitivity, 90% specificity), whereas Heller Murray et al. [6] reported lower cut-off values of −1.92 ST (VH subgroups vs. controls), though the sensitivity of these values were low, ranging from 29–43%.

Onset cycle 1 trended lower in individuals with voice disorders but showed substantial overlap with the range of values reported by those with typical voices. For example, studies reporting on adductor laryngeal dystonia hovered around 2.60 ST [2,4] and those with PD ranged from 1.75 ST [5] to 2.70 ST [49]. Studies on patients with VH were similar, reporting a range of 1.18 to 2.54 ST, with a median of 1.90 ST. ROC analyses showed close-to-chance detection rates [6,27], indicating a poor ability to discriminate between those with and without VH.

3.4. Between- and Within-Subject Analyses

A total of 12 studies completed formalized comparisons between experimental and control groups, with 9 studies finding statistically significant differences. The strongest evidence was found in comparisons involving adults with typical voices and subjects with VH, wherein 100% of studies (*n* = 4) reported significantly lower RFF values for those with VH. However, discrepancies arose when trying to pinpoint whether offset or onset cycles were different. One study reported significantly lower offsets [27], one reported significantly lower onsets [47], and the remaining two studies reported lower RFF values across both offsets and onsets [6,13]. Across other studies, there was consistent evidence

that patients with PD and those with adductor laryngeal dystonia had significantly lower RFF values as a whole compared to those with typical voices [4,5,49].

Other studies examined the impact of RFF across the lifespan. One study compared children who do and do not stutter [30] and another compared children with and without vocal fold nodules [35]; neither found differences between groups for any RFF measure. However, Heller Murray et al. [35] reported that RFF onset cycle 1 was lower in younger children compared to older children, pointing to a potential maturational effect for revoicing. A comparison undertaken between young children (age 4), older children (age 8), and young adults showed similar results, wherein young children had lower onset cycles (1 and 2) compared to young adults [46]. A study by Park et al. [44] examined differences between young adults and aging adults, finding no differences in RFF values; however, a study by Watson [12] did find significant differences for onset values only.

Within-subject analyses were undertaken to understand RFF sensitivity to subject fatigue. Four studies examined how vocal load impacted RFF [32,34,36,43], including short-term (a few hours) and long-term (multiple weeks). Findings were mixed: pre- to post-fatigue comparisons showed that offset cycle 10 may assist in tracking fatigue over the workday [43] but onset cycle 1 seemed to be more sensitive to long-term tracking over several weeks [34]. Fujiki et al. [32] saw no impact of a lab-based vocal loading task on offset cycle 10 or onset cycle 1, whereas Kagan and Heaton [36] reported composite changes in both measures following a formalized loading paradigm.

The relationship between the proposed physiological mechanisms of RFF (i.e., kinematics, laryngeal muscle tension, and aerodynamics) and RFF offset cycle 10 and onset cycle 1 were examined via kinematic estimates of laryngeal stiffness [1,44], an aerodynamic ratio of vocal intensity to estimated subglottal pressure [7], and vocal fold ab/adduction degree and timing [44]. To create within-subject variation, speakers with typical voices modulated vocal effort and/or strain across several voice productions. Group analyses yielded consistently poor-to-moderate relationships between RFF and the physiological parameters of laryngeal stiffness, aerodynamic ratio, and both kinematic degree and timing. Importantly, however, within-subject analyses demonstrated a marked improvement in the strength of these relationships, ranging from moderate to strong [1,7]. Accordingly, the authors supported the use of RFF as a clinical tracking tool to monitor an individual's voice changes over a course of therapeutic intervention (where it was assumed that vocal effort and/or strain would decrease over time).

Four studies specifically investigated how RFF may change over the course of various interventions. Two studies examined RFF changes over the course of vocal therapy, including a single-session treatment paradigm as well as a full course involving several therapy sessions. Roy et al. [47] found that onset values became higher and more similar to control values in a group of 111 patients with MTD following one hour of manual circumlaryngeal therapy. However, there was no effect of therapy on offset cycle values. Similarly, a group of 16 women (3 with nodules, 13 with MTD) completed a successful course of voice therapy, which resulted in significant improvements in RFF cycles as a whole [24]. The largest change was in onset cycle 1, with post-therapy values (M = 2.71 ST) being significantly greater than pre-therapy values (M = 1.90 ST). Interestingly, a comparison of individuals with phonotraumatic lesions pre-/post-surgical removal resulted in no significant changes to RFF [13]. The authors hypothesized that the surgical intervention only acted to change laryngeal anatomy, but did not mitigate the hyperfunctional behaviors and associated muscular tension, leading to no change in RFF values. The final treatment study investigated how low-level light therapy may promote healing and reduce inflammation in adults with typical voices undergoing a vocal fatigue paradigm [36]. Subjects were divided into four intervention groups: infrared wavelength, red wavelength, heat, and no-heat light (control). Due to the small sample size (n = 4 for each group), a formalized statistical analysis could not be undertaken; however, the red-light group demonstrated a faster return-to-baseline trend in RFF values when compared to all other groups.

3.5. Perceptual Measures: Effort, Strain, and Dysphonia

The relationship between RFF values and perceptual measures of vocal effort, strain, and dysphonia have varied widely across studies (*n* = 7). Lien et al. [7] reported poor relationships between listener-perception of vocal effort and offset cycle 10 and onset cycle 1 (R^2 = 0.21 and R^2 = 0.26, respectively) in individuals with typical voices while self-modulating their vocal effort. The relationship strength increased to moderate for offset cycle 10 (R^2 = 0.46) and onset cycle 1 (R^2 = 0.56) when analyzed within speaker. Conversely, relationships were significant between offset cycle 10 and listener-perceived ratings of vocal effort when analyzed by group across three other studies in those with typical and disordered voices, but no relationships were found for onset cycle 1 [4,8,27].

Roy et al. [47] found that RFF onset cycle 1 values were significantly correlated with overall dysphonia severity in a large group (*n* = 111) of patients with MTD. However, Stepp and colleagues [27] also investigated subjects with VH and reported weak relationships between onset cycle 1 and dysphonia ratings. Buckley et al. [2] took a slightly different approach in a group of subjects with adductor laryngeal dystonia by examining a combination of several acoustic measures, finding that offset cycle 10, changes between specific RFF cycles (e.g., change from onset cycle 1–6) and additional spectral characteristics were significantly related to dysphonia in their model.

A study by Park et al. [9] manually manipulated cycles in acoustic waveforms to understand how raising and lowering RFF would affect auditory-perceptual ratings of vocal strain. Eight adults without voice disorders produced VCV utterances in their typical voices and then again with maximal vocal effort. When intervocalic cycle values were artificially reduced in the typical voice samples, listeners perceived the voices as having greater strain; conversely, when RFF values were artificially raised in voicing samples collected during effortful speech, they were perceived as having less strain. This manipulation paradigm is perhaps some of the most convincing evidence of how cycle changes within a speaker can contribute to changes in perceived vocal characteristics. The directionality of the changes (i.e., lower RFF increased perceived strain while higher RFF reduced perceived strain) was consistent with the existing theoretical model of RFF not only being an indicator of laryngeal tension, but one that may be perceived by listeners through strained speech.

4. Discussion

This scoping review summarized the current state of research on RFF as a tool for assessing and tracking changes in tension-based voice disorders. We found 37 articles published since 1998 that lend support to RFF as a useful metric to characterize the human voice. The identified studies were heterogenous in nature, spanning patient populations, speech stimuli, extraction methods, and statistical testing, among other factors. Despite this heterogeneity, these works form a clear picture as to the potential utility of RFF for clinical voice evaluation, as well as areas that need further investigation prior to the clinical adoption of RFF.

4.1. Method of RFF Computation

Prior work on RFF indicates that approximately 20–30 min of manual processing time is necessary to calculate a reliable estimate for a single subject. Numerous studies (e.g., [51]) discuss the clinical implications of such an extensive processing time (not to mention the considerable time required to train oneself to manually estimate RFF), which point to a need for faster, automated methodologies for RFF to be implemented in voice clinics. Additional review of the literature indicated that there was, indeed, an increase in algorithmic investigations and implementation of these methodologies since their availability in 2016; 9 of the 17 studies (53%) published since 2016 have utilized algorithmic extraction, indicating that algorithmic adoption is a viable option for many researchers.

The aRFF algorithm [39] is the most commonly used automated method to date. This algorithm, much like software used during manual estimation (e.g., Praat), uses autocorrela-

tion to determine cycle f_o. It follows that one would expect a close correspondence between automated and manual estimates. Lien et al. [39] reported a strong relationship ($r \geq 0.82$) between aRFF and manual estimates for offset and onset values; however, the authors also found that the aRFF algorithm consistently estimates higher offset cycle 10 values ($+0.22$ to $+0.41$ ST) and lower onset cycle 1 values (-0.11 to -0.10 ST) when compared to manual estimates. Although these systematic errors are smaller than the increase in RFF that can be expected after undergoing voice therapy [24] (indicating that, on average, clinically meaningful changes in RFF will not be masked by errors associated with using the RFF algorithm), it is still important to consider the source of algorithm errors. One reason for this discrepancy may be that manual estimates allow the user to visually identify voicing offset and onset, whereas the algorithm must use computational techniques for vowel-fricative boundary identification. Boundary cycle shifts are most likely to impact the cycle values closest to the fricative (i.e., offset cycle 10 and onset cycle 1) and require refinement for consistent value interpretation and clinical adoption.

Each new algorithm iteration has improved upon the last by increasing the accuracy of cycle detection through various signal processing considerations. For example, Vojtech et al. [51] incorporated new pitch estimation methods in aRFF-AP to more closely align algorithm extracted RFF values with manually acquired values. Within this algorithm, the authors also introduced a set of signal-specific parameters to account for differences in voice sample characteristics (e.g., overall severity of dysphonia, signal acquisition quality) based on the acoustic measure, pitch strength [55,56]. In doing so, errors affecting the accuracy of the RFF algorithm were reduced by 88.4% relative to manually acquired values.

Further exploration into improving the accuracy of algorithmically extracted RFF values has yielded a new set of acoustic feature processing techniques to locate the voicing cycles closest to the voiceless consonant (i.e., offset cycle 10, onset cycle 1). This most recent version of the RFF algorithm (aRFF-APH) [52] was built by integrating information from simultaneous recordings made using a microphone and high-speed flexible laryngoscopy. The authors were able to identify the physiological—rather than acoustic—initiation and termination of vocal fold vibration that marks the boundaries of the voiceless consonant. Interestingly, the study reported that both algorithm iterations (aRFF-AP, aRFF-APH) were significantly more likely to identify offset cycle 10 and onset cycle 1 values that coincided with physiologically identified vocal fold vibratory timings than manually acquired values, with aRFF-AP having the strongest correspondence with physiologic events. The results of this work call into question the use of manual RFF as a "ground truth" comparison for algorithmic testing. Instead, aligning RFF automated extraction with physiologic events may increase the precision of the measure and provide insight into the underlying physiological factors proposed to influence RFF values. Although the aRFF-APH algorithm demonstrated the highest association with physiologically identified vocal fold vibratory timings, the aRFF-AP algorithm remains the gold-standard method for algorithmically estimating RFF due to its validation across a large dataset (483 independent speakers) and broad range of vocal disorders and severities. A large-scale analysis is still needed with the aRFF-APH algorithm to determine its generalizability across various voice disorders.

Although algorithms developed for microphone signals (aRFF, aRFF-AP) are publicly available and relatively easy to use, automated RFF extraction continues to pose issues for clinical practice. First, these RFF algorithms are semi-automated, meaning that they continue to require user interaction to inform extraction methods and determine extraction accuracy. The algorithm interface allows the user to check the accuracy of the voiceless consonant detection from the acoustic waveform and override the automated decision, if necessary, thereby improving extraction accuracy, but also increasing the time needed for processing. Second, all publicly available algorithms [57] require the use of proprietary MATLAB software, including several toolboxes (e.g., econometrics, curvefitting, signal processing) at additional cost. MATLAB is not a common software used by most voice clinics, further limiting the use of algorithms for RFF processing. Ideally, automated RFF methods would expand into freely available software to increase its usability at clinical

sites. Third, and perhaps most importantly, automated algorithms can only extract RFF from VCV stimuli. An algorithm that could extract RFF from continuous speech would be helpful for retrospective analyses of pre-established voice databases that include standard reading passages (e.g., Rainbow Passage). Then, researchers could undertake a large-scale analysis across patients of various clinical diagnoses, which would allow for refinement of clinical ranges, cut-off values for classification between vocal health and disorder, and clinical tracking over the course of intervention.

A more recent algorithmic iteration expanded previous work from acoustic microphone signals to signals gathered with neck-surface accelerometers [33]. A neck-surface accelerometer is a small sensor (approximately the size of a dime) placed on the anterior surface of the neck, inferior to the location of the cricoid cartilage but superior to the sternal notch, at midline. The sensor captures vibration from the surface of the skin and has the benefit of reduced signal disruption from environmental/background noise and speech frication and/or aspiration noise [58]. Furthermore, the accelerometer does not adequately capture oral resonances, making it difficult-to-impossible to discern what the speaker is saying, thus protecting speaker privacy during communication exchanges.

It follows that accelerometers—coupled with smartphone technology—are under development for ambulatory voice monitoring and biofeedback for those at risk for voice disorders (e.g., occupational voice users) and those with current voice disorders [59–62]. The incorporation of RFF into daily biofeedback could be beneficial for within-subject clinical tracking and/or the identification of maladaptive tension-based vocal patterns in pre-clinical populations. Lien and colleagues [37,41] have shown that RFF calculated from accelerometer signals are lower when compared to those calculated from simultaneously acquired microphone signals. Therefore, accelerometer-derived RFF values need to be interpreted with caution and only compared to other values obtained from accelerometers. Finally, similar to the recommendation for microphone-based RFF algorithmic processing, extraction from continuous speech would increase the usability of accelerometer-based RFF as a monitoring tool during daily communication. These areas of future development are ongoing and provide a potential exciting avenue for incorporation of continuous monitoring to assist in the remediation of voice disorders, and the prevention of aberrant vocal behaviors as well.

4.2. Interpreting RFF in Clinical Practice

Regardless of the method used to calculate RFF, it remains that one of the primary challenges hindering the clinical adoption of RFF is a lack of standardized values for typical and disordered voices. Unfortunately, our review of the literature does not provide a clear answer as to which RFF values should be used for a clinically meaningful distinction of vocal health or disorder. On the one hand, we found robust evidence that individuals with voice disorders have consistently lower offset cycle 10 values when directly comparing the range of their reported mean values to those reported in adults with typical voices. Onset cycle 1, however, showed considerable overlap between typical and disordered groups. These findings are surprising since statistical comparisons between those with and without voice disorders were significant in 9 out of 12 studies. This discrepancy could be due to analyses that combined offset and onset cycles rather than reporting offset and onset values in isolation. Thus, at present, there seems to be no clear recommendation for exactly which cycle, or combination of cycles, would effectively distinguish between typical and disordered voices. This area of research is still ongoing, with recent studies looking to enhance the sensitivity of RFF via advancing computational methods (e.g., automated, rule-based decision criteria [51]) as well as combining RFF with other acoustic parameters to improve the accuracy of RFF as a clinical tool of diagnosis [63].

Importantly, offset cycle 10 and onset cycle 1 ranges reported in this review were created without consideration of the number of subjects enrolled across studies, leading to an equal weighting for studies that enrolled few subjects and those that enrolled many. Instead, a weighted average would be the best way to help establish normative ranges

and values across healthy and disordered populations. At present, the number of studies reporting RFF values for adults with typical voices ($n = 14$) may be sufficient to establish an acceptable normative range, but more studies need to report raw RFF values (*M*, SD, range) for aging adults and children with typical voices to establish normative ranges across the lifespan.

At present, there seems to be stronger evidence for the utility of RFF as a clinical tracking tool for within-subject changes over several voice productions as well as for quantifying progress pre- and post-therapy. Tracking of physiological parameters (kinematic stiffness, aerodynamics) were shown to be moderate-to-strongly correlated with RFF for individual analyses of vocally healthy speakers varying their vocal effort and strain [1,7]. Furthermore, Roy et al. [47] and Stepp et al. [24] reported improvements in onset cycle 1 values (i.e., more positive values) in patients with tension-based voice disorders following vocal therapy. Despite the promise of these results, there were few studies that examined RFF as a clinical outcome measure [13,24,36,47].

Indeed, a systematic review with meta-analysis is considered the highest level of evidence for understanding treatment efficacy and establishing clinical guidelines [64]. However, different intervention approaches (single therapy visit, multiple therapy visits, surgery) and various treatment groups (e.g., MTD, VH, polyps, typical) across the studies described in our review, make a meta-analysis impossible to undertake at this time. We recommend that future treatment studies of patients with tension-based voice disorders incorporate RFF into their standard voice protocols to increase the number of studies available for a large-scale analysis and assist in validating RFF as a treatment outcome. This would not only provide evidence for RFF as an appropriate outcome measure in specific patient populations, but also help to define what a clinically meaningful change in RFF may be over the course of intervention.

A second recommendation would be to incorporate RFF measures into the therapeutic assessment protocols and tracking for patients with listener-perceived vocal strain (defined as excessive vocal effort [65]) or self-reported vocal effort. Historically, vocal strain and effort have proved difficult to quantify via acoustic measures [66–68]. This may be due to the variation of strained and effortful productions across speakers, as well as the potential overlap of percepts in individuals with voice disorders, in which a person with a strained voice can also present with a breathy voice. Our review of the literature indicates that RFF may be an appropriate clinical tool related to the percepts of strain and effort with offset cycle 10, in particular, being significantly correlated with either speaker or listener-ratings [4,7,8,27].

Perhaps the most substantial evidence for RFF as an objective indicator of vocal strain was described in Park et al. [9]. The study design controlled for the problems of speaker variation and overlapping voicing percepts by synthetically altering RFF cycle values in speech samples gathered from those with typical voices. The authors manipulated RFF cycle values in two ways: (i) artificially lowering RFF cycle values during typical speech, which was hypothesized to increase the perception of strain due to the lower values seen in individuals with tension-based voice disorders, and (ii) artificially raising RFF cycle values in productions of maximal vocal effort so that the cycle values would resemble those found in typical voice productions. The hypotheses were supported when these manipulations resulted in changes to perceived strain in the speakers, wherein lowering RFF values acted to increase the perception of strain, and increasing RFF values reduced perceived strain. The authors incorporated an additional condition into the study in which they artificially increased mid-to-high frequency noise in the signal (which they quantified via a harmonics-to-noise ratio) and, subsequently, saw an increase in perceived strain. This study provides a foundation for future perceptual work on RFF, in that it may prove beneficial to study RFF in combination with other acoustic measures [8] in order to increase the correspondence between acoustic and perceptual measures.

Finally, a thorough understanding of the relationship between RFF and underlying laryngeal physiology is needed to appropriately interpret RFF values. That is, in order

to use RFF as a clinical outcome measure, it should be validated as an indicator of laryngeal tension. In our review, few studies investigated the laryngeal factors proposed by Stepp et al. [24]; however, the results of these studies provide preliminary evidence for RFF being related to measures of laryngeal stiffness [1,44], aerodynamics [7], and vocal fold adduction parameters [44]. Our review did not delve into the physiological basis of RFF, partially due to the fact that we limited our search to studies that enrolled human subjects, excluding modeling and theoretical studies. Recent work on modeling voicing offset behaviors [69] suggests a physiological relationship between abduction duration and RFF values; however, these findings should be expanded upon, and evaluated with empirical testing in human subjects. Combining computational modeling strategies with human subject research is an important step towards isolating the differential components associated with intervocalic voicing offset and onset (i.e., laryngeal kinematics, aerodynamics, and tension) and, in turn, may provide useful insights into the clinical utility of RFF for assessing and treating vocal disorders.

4.3. Limitations

Although every effort was made to make this scoping review as current as possible, researchers have continued to publish on RFF. Since the hand search completed in September of 2021, there are at least five more published (or in press) studies that may have qualified for this review [63,70–73]. We acknowledge that all reviews are meant to be updated and expanded upon, and recommend these studies be considered in future reviews on RFF.

This scoping review focused on specific topics pertaining to the clinical adoption of RFF in patient populations and the processing methods required to do so. Still, there are other topics that were not extensively discussed in our review, but should be topics of future summaries. These include (but are not limited to): (i) the impacts of phoneme, stress, and pitch on RFF measures, (ii) comparisons between VCV and continuous speech estimations of RFF values, (iii) discussions on the inability to calculate RFF in specific cases (e.g., glottalization, fewer than 10 voicing cycles, no established steady state) and the challenges of missing data in individuals with voice disorders, and (iv) physiological modeling [69] and cross-methodological studies to validate RFF as a measure of tension, aerodynamics, and/or laryngeal kinematics [73]. This last point, in particular, is needed to provide quantitative evidence to support, refute, or update the models proposed by Stepp et al. [24] and Heller Murray et al. [6].

5. Conclusions

RFF is a relatively new voice measure that has been described in the literature since 1998. At present, the clinical adoption of RFF is limited by two primary factors, including the availability of automated processing methods, and the interpretation of RFF values. Our review showed that automated algorithms have been incorporated into nearly 55% of all research studies since their availability in 2016. The algorithms have undergone several iterations and have been extended into ambulatory monitoring of voice disorders, a key clinical advancement for daily monitoring and biofeedback for patients. We found evidence for RFF as a potential treatment outcome measure and clinical tracking tool. However, more intervention studies with RFF as a clinical outcome measure are needed in order to complete a formalized meta-analysis and identify changes in RFF that are clinically meaningful and applicable. Finally, investigations into the underlying physiology of RFF are still needed to fully understand how this measure corresponds to the identified factors of tension, aerodynamics, and kinematics.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app12168121/s1, Table S1: Studies focused on automated algorithmic processing of RFF; Table S2: Summary of studies (*n* = 32) examining RFF in people with and without voice disorders.

Appendix A. Example Search String: Embase

(((('relative fundamental frequenc*':ti,ab OR 'rff':ti,ab OR 'fundamental frequency'/exp) OR (('fundamental frequenc*':ti,ab OR 'f0':ti,ab) AND 'onset':ti,ab) OR (('fundamental frequenc*':ti,ab OR 'f0':ti,ab) AND 'offset':ti,ab) OR (('fundamental frequenc*':ti,ab OR 'f0':ti,ab) AND 'cycle to cycle':ti,ab) OR (('fundamental frequenc*':ti,ab OR 'f0':ti,ab) AND ('voice'/exp OR 'voice*':ti,ab OR 'vocal*':ti,ab))) AND ('speech'/exp OR 'speech':ti,ab OR 'voice'/exp OR 'voice*':ti,ab OR 'vocal*':ti,ab)) NOT ('animal'/de OR 'animal experiment'/de OR 'nonhuman'/de).

References

1. McKenna, V.S.; Heller Murray, E.S.; Lien, Y.-A.S.; Stepp, C.E. The Relationship Between Relative Fundamental Frequency and a Kinematic Estimate of Laryngeal Stiffness in Healthy Adults. *J. Speech Lang. Hear. Res.* **2016**, *59*, 1283–1294. [CrossRef] [PubMed]
2. Buckley, D.P.; Cadiz, M.D.; Eadie, T.L.; Stepp, C.E. Acoustic Model of Perceived Overall Severity of Dysphonia in Adductor-Type Laryngeal Dystonia. *J. Speech Lang. Hear. Res.* **2020**, *63*, 2713–2722. [CrossRef] [PubMed]
3. Ramig, L.O.; Verdolini, K. Treatment Efficacy: Voice Disorders. *J. Speech Lang Hear. Res.* **1998**, *41*, S101–S116. [CrossRef] [PubMed]
4. Eadie, T.L.; Stepp, C.E. Acoustic Correlate of Vocal Effort in Spasmodic Dysphonia. *Ann. Otol. Rhinol. Laryngol.* **2013**, *122*, 169–176. [CrossRef]
5. Goberman, A.M.; Blomgren, M. Fundamental Frequency Change During Offset and Onset of Voicing in Individuals with Parkinson Disease. *J. Voice* **2008**, *22*, 178–191. [CrossRef]
6. Heller Murray, E.S.; Lien, Y.-A.S.; Van Stan, J.H.; Mehta, D.D.; Hillman, R.E.; Noordzij, J.P.; Stepp, C.E. Relative Fundamental Frequency Distinguishes Between Phonotraumatic and Non-Phonotraumatic Vocal Hyperfunction. *J. Speech Lang. Hear. Res.* **2017**, *60*, 1507–1515. [CrossRef]
7. Lien, Y.-A.S.; Michener, C.M.; Eadie, T.L.; Stepp, C.E. Individual Monitoring of Vocal Effort with Relative Fundamental Frequency: Relationships with Aerodynamics and Listener Perception. *J. Speech Lang. Hear. Res.* **2015**, *58*, 566–575. [CrossRef]
8. McKenna, V.S.; Stepp, C.E. The relationship between acoustical and perceptual measures of vocal effort. *J. Acoust. Soc. Am.* **2018**, *144*, 1643–1658. [CrossRef]
9. Park, Y.; Cádiz, M.D.; Nagle, K.F.; Stepp, C.E. Perceptual and Acoustic Assessment of Strain Using Synthetically Modified Voice Samples. *J. Speech Lang. Hear. Res.* **2020**, *63*, 3897–3908. [CrossRef]
10. Park, Y.; Stepp, C.E. The Effects of Stress Type, Vowel Identity, Baseline f0, and Loudness on the Relative Fundamental Frequency of Individuals with Healthy Voices. *J. Voice* **2019**, *33*, 603–610. [CrossRef]
11. Fairbanks, G. *Voice and Articulation Drillbook*, 2nd ed.; Harper & Row: New York, NY, USA, 1960.
12. Watson, B.C. Fundamental frequency during phonetically governed devoicing in normal young and aged speakers. *J. Acoust. Soc. Am.* **1998**, *103*, 3642–3647. [CrossRef]
13. Stepp, C.E.; Hillman, R.E.; Heaton, J.T. The Impact of Vocal Hyperfunction on Relative Fundamental Frequency During Voicing Offset and Onset. *J. Speech Lang. Hear. Res.* **2010**, *53*, 1220–1226. [CrossRef]
14. Halle, M.; Stevens, K. *A note on Laryngeal Features*; Quartely Prog. Rep. 101; Research Laboratory of Electronics MIT: Cambridge, MA, USA, 1971; pp. 198–212.
15. Rothenberg, M.; Mahshie, J.J. Monitoring Vocal Fold Abduction through Vocal Fold Contact Area. *J. Speech Lang. Hear. Res.* **1988**, *31*, 338–351. [CrossRef]

16. Löfqvist, A.; Baer, T.; McGarr, N.S.; Story, R.S. The cricothyroid muscle in voicing control. *J. Acoust. Soc. Am.* **1989**, *85*, 1314–1321. [CrossRef]

17. Hirano, M. Morphological Structure of the Vocal Cord as a Vibrator and its Variations. *Folia Phoniatr. Logop.* **1974**, *26*, 89–94. [CrossRef]

18. Kent, R.D.; Moll, K.L. Vocal-Tract Characteristics of the Stop Cognates. *J. Acoust. Soc. Am.* **1969**, *46*, 1549–1555. [CrossRef]

19. House, A.S.; Fairbanks, G. The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels. *J. Acoust. Soc. Am.* **1953**, *25*, 105–113. [CrossRef]

20. Löfqvist, A.; Koenig, L.L.; Mcgowan, R.S. Vocal tract aerodynamics in /aCa/ utterances: Measurements. *Speech Commun.* **1995**, *16*, 49–66. [CrossRef]

21. Hirano, M.; Ohala, J.; Vennard, W. The Function of Laryngeal Muscles in Regulating Fundamental Frequency and Intensity of Phonation. *J. Speech Hear. Res.* **1969**, *12*, 616–628. [CrossRef]

22. Ohde, R.N. Fundamental frequency as an acoustic correlate of stop consonant voicing. *J. Acoust. Soc. Am.* **1984**, *75*, 224–230. [CrossRef]

23. Ohde, R.N. Fundamental frequency correlates of stop consonant voicing and vowel quality in the speech of preadoles-cent children. *J. Acoust. Soc. Am.* **1985**, *78*, 1554–1561. [CrossRef]

24. Stepp, C.E.; Merchant, G.R.; Heaton, J.T.; Hillman, R.E. Effects of Voice Therapy on Relative Fundamental Frequency During Voicing Offset and Onset in Patients with Vocal Hyperfunction. *J. Speech Lang. Hear. Res.* **2011**, *54*, 1260–1266. [CrossRef]

25. Hillman, R.E.; Holmberg, E.B.; Perkell, J.S.; Walsh, M.; Vaughan, C. Objective assessment of vocal hyperfunction: An experi-mental framework and initial results. *J. Speech Hear. Res.* **1989**, *32*, 373–392. [CrossRef]

26. Patel, R.R.; Awan, S.N.; Barkmeier-Kraemer, J.; Courey, M.; Deliyski, D.; Eadie, T.; Paul, D.; Švec, J.G.; Hillman, R. Recommended Protocols for Instrumental Assessment of Voice: American Speech-Language-Hearing Association Expert Panel to Develop a Protocol for Instrumental Assessment of Vocal Function. *AJSLP* **2018**, *27*, 887–905. [CrossRef]

27. Stepp, C.E.; Sawin, D.E.; Eadie, T.L. The Relationship Between Perception of Vocal Effort and Relative Fundamental Frequency During Voicing Offset and Onset. *J. Speech Lang. Hear. Res.* **2012**, *55*, 1887–1896. [CrossRef]

28. Tricco, A.C.; Lillie, E.; Zarin, W.; O'Brien, K.K.; Colquhoun, H.; Levac, D.; Moher, D.; Peters, M.D.J.; Horsley, T.; Straus, S.E.; et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann. Intern. Med.* **2018**, *169*, 467–473. [CrossRef]

29. Ouzzani, M.; Hammady, H.; Fedorowicz, Z.; Elmagarmid, A. Rayyan—A web and mobile app for systematic reviews. *Syst. Rev.* **2016**, *5*, 210. [CrossRef]

30. Arenas, R.M.; Zebrowski, P.M.; Moon, J.B. Phonetically governed voicing onset and offset in preschool children who stutter. *J. Fluen. Disord.* **2012**, *37*, 179–187. [CrossRef]

31. Dahl, K.L.; Stepp, C.E. Changes in Relative Fundamental Frequency Under Increased Cognitive Load in Individuals with Healthy Voices. *J. Speech Lang. Hear. Res.* **2021**, *64*, 1189–1196. [CrossRef]

32. Fujiki, R.B.; Chapleau, A.; Sundarrajan, A.; McKenna, V.; Sivasankar, M.P. The Interaction of Surface Hydration and Vocal Loading on Voice Measures. *J. Voice* **2017**, *31*, 211–217. [CrossRef]

33. Groll, M.D.; Vojtech, J.M.; Hablani, S.; Mehta, D.D.; Buckley, D.P.; Noordzij, J.P.; Stepp, C. Automated Relative Fundamental Frequency Algorithms for Use with Neck-Surface Accelerometer Signals. *J Voice* **2022**, *36*, 156–169. [CrossRef] [PubMed]

34. Heller Murray, E.S.; Hands, G.L.; Calabrese, C.R.; Stepp, C.E. Effects of Adventitious Acute Vocal Trauma: Relative Fun-damental Frequency and Listener Perception. *J. Voice* **2016**, *30*, 177–185. [CrossRef] [PubMed]

35. Heller Murray, E.S.; Segina, R.K.; Woodnorth, G.H.; Stepp, C.E. Relative Fundamental Frequency in Children with and without Vocal Fold Nodules. *J. Speech Lang. Hear. Res.* **2020**, *63*, 361–371. [CrossRef] [PubMed]

36. Kagan, L.S.; Heaton, J.T. The Effectiveness of Low-Level Light Therapy in Attenuating Vocal Fatigue. *J. Voice* **2017**, *31*, 384.e15–384.e23. [CrossRef]

37. Lien, Y.-A.S.; Calabrese, C.R.; Michener, C.M.; Heller Murray, E.; Van Stan, J.H.; Mehta, D.D.; Hillman, R.E.; Noordzij, J.P.; Stepp, C.E. Voice Relative Fundamental Frequency Via Neck-Skin Acceleration in Individuals with Voice Disorders. *J. Speech Lang. Hear. Res.* **2015**, *58*, 1482–1487. [CrossRef]

38. Lien, Y.-A.S.; Gattuccio, C.I.; Stepp, C.E. Effects of Phonetic Context on Relative Fundamental Frequency. *J. Speech Lang. Hear. Res.* **2014**, *57*, 1259–1267. [CrossRef]

39. Lien, Y.-A.S.; Heller Murray, E.S.; Calabrese, C.R.; Michener, C.M.; Van Stan, J.H.; Mehta, D.D.; Hillman, R.E.; Noordzij, J.P.; Stepp, C.E. Validation of an Algorithm for Semi-automated Estimation of Voice Relative Fundamental Frequency. *Ann. Otol. Rhinol. Laryngol.* **2017**, *126*, 712–716. [CrossRef]

40. Lien, Y.-A.S.; Stepp, C.E. Automated estimation of relative fundamental frequency 2013. In Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), Osaka, Japan, 3–7 July 2013; pp. 2136–2139.

41. Lien, Y.-A.S.; Stepp, C.E. Comparison of voice relative fundamental frequency estimates derived from an accelerometer signal and low-pass filtered and unprocessed microphone signals. *J. Acoust. Soc. Am.* **2014**, *135*, 2977–2985. [CrossRef]

42. McKenna, V.S.; Kendall, C.L.; Patel, T.H.; Howell, R.J.; Gustin, R.L. Impact of Face Masks on Speech Acoustics and Vocal Effort in Healthcare Professionals. *Laryngoscope* **2022**, *132*, 391–397. [CrossRef]

43. McKenna, V.S.; Patel, T.H.; Kendall, C.L.; Howell, R.J.; Gustin, R.L. Voice Acoustics and Vocal Effort in Mask-Wearing Healthcare Professionals: A Comparison Pre- and Post-Workday. *J. Voice* **2021**, S089219972100151X. [CrossRef]

44. Park, Y.; Wang, F.; Díaz-Cádiz, M.; Vojtech, J.M.; Groll, M.D.; Stepp, C.E. Vocal fold kinematics and relative fundamental frequency as a function of obstruent type and speaker age. *J. Acoust. Soc. Am.* **2021**, *149*, 2189–2199. [CrossRef]

45. Park, Y.; Stepp, C.E. Test–Retest Reliability of Relative Fundamental Frequency and Conventional Acoustic, Aerodynamic, and Perceptual Measures in Individuals with Healthy Voices. *J. Speech Lang. Hear. Res.* **2019**, *62*, 1707–1718. [CrossRef]

46. Robb, M.P.; Smith, A.B. Fundamental frequency onset and offset behavior: A comparative study of children and adults. *J. Speech Lang. Hear. Res.* **2002**, *45*, 446–456. [CrossRef]

47. Roy, N.; Fetrow, R.A.; Merrill, R.M.; Dromey, C. Exploring the Clinical Utility of Relative Fundamental Frequency as an Objective Measure of Vocal Hyperfunction. *J. Speech Lang. Hear. Res.* **2016**, *59*, 1002–1017. [CrossRef]

48. Smith, A.B.; Robb, M.P. Factors underlying short-term fundamental frequency variation during vocal onset and offset. *Speech Lang. Hear.* **2013**, *16*, 208–214. [CrossRef]

49. Stepp, C.E. Relative fundamental frequency during vocal onset and offset in older speakers with and without Parkinson's disease. *J. Acoust. Soc. Am.* **2013**, *133*, 1637–1643. [CrossRef]

50. Van Mersbergen, M.; Lanza, E. Modulation of Relative Fundamental Frequency During Transient Emotional States. *J. Voice* **2019**, *33*, 894–899. [CrossRef]

51. Vojtech, J.M.; Segina, R.K.; Buckley, D.; Kolin, K.R.; Tardif, M.C.; Noordzij, J.P.; Stepp, C.E. Refining algorithmic estimation of relative fundamental frequency: Accounting for sample characteristics and fundamental frequency estimation method. *J. Acoust. Soc. Am.* **2019**, *146*, 3184–3202. [CrossRef]

52. Vojtech, J.; Cilento, D.; Luong, A.; Noordzij, J.; Diaz-Cadiz, M.; Groll, M.; Buckley, D.; McKenna, V.; Noordzij, J.; Stepp, C. Acoustic Identification of the Voicing Boundary during Intervocalic Offsets and Onsets Based on Vocal Fold Vibratory Measures. *Appl. Sci.* **2021**, *11*, 3816. [CrossRef]

53. Weerathunge, H.R.; Segina, R.K.; Tracy, L.; Stepp, C.E. Accuracy of Acoustic Measures of Voice via Telepractice Videocon-ferencing Platforms. *J. Speech Lang. Hear. Res.* **2021**, *64*, 2586–2599. [CrossRef]

54. Boersma, P. Praat, a system for doing phonetics by computer. *Glot. Int.* **2001**, *5*, 341–345.

55. Kopf, L.M.; Jackson-Menaldi, C.; Rubin, A.D.; Skeffington, J.; Hunter, E.J.; Skowronski, M.D.; Shrivastav, R. Pitch Strength as an Outcome Measure for Treatment of Dysphonia. *J. Voice* **2017**, *31*, 691–696. [CrossRef]

56. Shrivastav, R.; Eddins, D.A.; Anand, S. Pitch strength of normal and dysphonic voices. *J. Acoust. Soc. Am.* **2012**, *131*, 2261–2269. [CrossRef]

57. Stepp, C.E. RFF. BU STEPP Lab Sensorimotor Rehabil Eng n.d. Available online: https://sites.bu.edu/stepplab/research/rff/ (accessed on 2 June 2022).

58. Cheyne, H.A.; Hanson, H.M.; Genereux, R.P.; Stevens, K.N.; Hillman, R.E. Development and Testing of a Portable Vocal Ac-cumulator. *J. Speech Lang. Hear. Res.* **2003**, *46*, 1457–1467. [CrossRef]

59. Hillman, R.E.; Heaton, J.T.; Masaki, A.; Zeitels, S.M.; Cheyne, H.A. Ambulatory Monitoring of Disordered Voices. *Ann. Otol. Rhinol. Laryngol.* **2006**, *115*, 795–801. [CrossRef]

60. Mehta, D.D.; Van Stan, J.H.; Zañartu, M.; Ghassemi, M.; Guttag, J.V.; Espinoza, V.M.; Cortés, J.P.; Cheyne, H.A.I.; Hillman, R.E. Using Ambulatory Voice Monitoring to Investigate Common Voice Disorders: Research Update. *Front. Bioeng. Biotechnol.* **2015**, *3*, 155. [CrossRef]

61. Van Stan, J.H.; Mehta, D.D.; Zeitels, S.M.; Burns, J.A.; Barbu, A.M.; Hillman, R.E. Average Ambulatory Measures of Sound Pressure Level, Fundamental Frequency, and Vocal Dose Do Not Differ Between Adult Females with Phonotraumatic Lesions and Matched Control Subjects. *Ann. Otol. Rhinol. Laryngol.* **2015**, *124*, 864–874. [CrossRef]

62. Van Stan, J.H.; Ortiz, A.J.; Sternad, D.; Mehta, D.D.; Huo, C.; Hillman, R.E. Ambulatory Voice Biofeedback: Acquisition and Retention of Modified Daily Voice Use in Patients with Phonotraumatic Vocal Hyperfunction. *Am. J. Speech Lang. Pathol.* **2022**, *31*, 409–418. [CrossRef]

63. Kapsner-Smith, M.R.; Díaz-Cádiz, M.E.; Vojtech, J.M.; Buckley, D.P.; Mehta, D.D.; Hillman, R.E.; Tracy, L.F.; Noordzij, J.P.; Eadle, T.L.; Stepp, C.E. Clinical Cutoff Scores for Acoustic Indices of Vocal Hyperfunction That Combine Relative Fundamental Frequency and Cepstral Peak Prominence. *J. Speech Lang. Hear. Res.* **2022**, *65*, 1349–1369. [CrossRef]

64. Tenny, S.; Varacallo, M. *Evidence Based Medicine*; StatPearls Publishing: Treasure Island, FL, USA, 2022.

65. Kempster, G.B.; Gerratt, B.; Abbott, K.V.; Barkmeier-Kraemer, J.; Hillman, R.E. Consensus Auditory-Perceptual Evaluation of Voice: Development of a Standardized Clinical Protocol. *Am. J. Speech Lang. Pathol.* **2009**, *18*, 124–132. [CrossRef]

66. Anand, S.; Bottalico, P.; Gray, C. Vocal Fatigue in Prospective Vocal Professionals. *J. Voice* **2019**, *35*, 247–258. [CrossRef] [PubMed]

67. Lowell, S.Y.; Kelley, R.T.; Awan, S.N.; Colton, R.H.; Chan, N.H. Spectral- and Cepstral-Based Acoustic Features of Dysphonic, Strained Voice Quality. *Ann. Otol. Rhinol. Laryngol.* **2012**, *121*, 539–548. [CrossRef] [PubMed]

68. Rosenthal, A.L.; Lowell, S.Y.; Colton, R.H. Aerodynamic and Acoustic Features of Vocal Effort. *J. Voice* **2014**, *28*, 144–153. [CrossRef] [PubMed]

69. Serry, M.A.; Stepp, C.E.; Peterson, S.D. Physics of phonation offset: Towards understanding relative fundamental frequency observations. *J. Acoust. Soc. Am.* **2021**, *149*, 3654–3664. [CrossRef]

70. Buckley, D.P.; Vojtech, J.M.; Stepp, C.E. Relative Fundamental Frequency in Individuals with Globus Syndrome and Muscle Tension Dysphagia. *J. Voice* **2021**, S0892199721003477. [CrossRef]

71. Dahl, K.L.; François, F.A.; Buckley, D.P.; Stepp, C.E. Voice and Speech Changes in Transmasculine Individuals Following Circumlaryngeal Massage and Laryngeal Reposturing. *Am. J. Speech Lang. Pathol.* **2022**, *31*, 1368–1382. [CrossRef]
72. McKenna, V.S.; Gustin, R.L.; Howell, R.J.; Patel, T.H.; Emery, M.B.; Kendall, C.L.; Kelliher, N.J. Developing Educational Health Modules to Improve Vocal Wellness in Mask-Wearing Occupational Voice Users. *J. Voice* **2021**, S0892199721003921. [CrossRef]
73. Vojtech, J.M.; Stepp, C.E. Effects of Age and Parkinson's Disease on the Relationship between Vocal Fold Abductory Kinematics and Relative Fundamental Frequency. *J. Voice* **2022**, S0892199722000704. [CrossRef]

Table S1. Studies focused on automated algorithmic processing of RFF.

Study	Study Population						Stimuli	Recording Device(s)	Study Summary and Findings
	Typical Voices			Disordered Voices					
	N	Sex	Age (Years) *M*±SD [range]	N	Sex	Age (Years) *M*±SD [range]			
Lien & Stepp [40]	12	6 M 6 F	20.0±2.0 [NR]	12 PD	7 M 5 F	69.0±11.0 [41,82]	VCV	Microphone	First iteration of an RFF algorithm, showing that algorithmic RFF estimates were similar to manual estimates, especially for voicing offset cycles and for individuals with PD. No statistical analysis was conducted to discern significant differences between manual and algorithmic RFF values.
Lien et al. [39]	36	9 M 27 F	41.0±19.0 [18,85]	154 (54 NPVH 81 PVH 19 Other)	38 M 116 F	41.0±17.0 [18,83]	VCV	Microphone	Large-scale study that sought to statistically examine algorithmic and manual RFF estimates. Semi-automated RFF estimates were highly correlated (*r* ≥ .82) with manual RFF estimates. Algorithm-derived offset cycle 10 values were slightly lower than manual estimates (−0.31 ST on average), whereas onset cycle 1 values were slightly higher (+0.10 ST on average). This algorithm used autocorrelation as its primary pitch detection method and is known as aRFF.
Vojtech et al. [51]	256	104 M 152 F	37.6±22.3 [18,100]	227 VD[a]	79 M 148 F	52.9±17.7 [18,84]	VCV	Microphone	Large-scale study that sought to increase the accuracy and precision of semi-automated RFF estimates. Different pitch estimation methods (e.g., Auditory-SWIPE′, Halcyon) were introduced into the algorithm and resulting values were compared to manual estimates. The acoustic measure, pitch strength, was used to account for differences in voice sample characteristics (dysphonia severity, signal quality). Results showed that the Auditory-SWIPE′ algorithm, in combination with pitch strength-based algorithm parameters, reduced errors between algorithmic and manual computed RFF values. This version of the algorithm is referred to as aRFF-AP.
Vojtech et al. [52]	69	36 M 33 F	43.2±23.1 [18,91]	53 (25 PD 20 NPVH 8 PVH)	24 M 29 F	49.5±18.4 [19,75]	VCV	Microphone, High-speed laryngoscopy	Exploratory study involving simultaneous recordings made using a microphone and high-speed laryngoscopy to compare acoustic and physiological voicing offset and onset boundaries. A new set of acoustic parameters were identified to tune the RFF algorithm to the known physiological boundaries, finding that the addition of these features significantly improved the algorithm from a previous version as well as from manual extraction. The new algorithm, called aRFF-APH, requires further validation with larger training and testing sets.
Groll et al. [33]	202	70 M 132 F	35.4±21.4 [18,100]	192 VD[b]	61 M 131 F	51.9±17.6 [18,84]	VCV	Microphone, Accelerometer	Details fully automated RFF algorithm developed for use with neck-surface accelerometer signals. RFF values were compared between the accelerometer algorithm, microphone-based algorithm, and manual extraction. Accelerometer-based values demonstrated larger errors than microphone-based values when compared to manual estimates. The authors surmised that this was due to differences between signal types rather than the algorithmic processing methods.

Note. *M*=mean; SD=standard deviation; M=male; F=female; NR=not reported; VD = voice disordered; PD = Parkinson's disease; PVH=phonotraumatic vocal hyperfunction (e.g., nodules, polyps, scar); NPVH=non-phonotraumatic vocal hyperfunction (e.g., muscle tension dysphonia); [a]=NPVH, PVH, PD, vocal fold paralysis, papilloma, globus sensation or dysphagia, infection, reflux, vocal fold atrophy. [b]=PD, NPVH, PVH, laryngeal dystonia.

Table S2. Summary of studies (*n* = 32) examining RFF in people with and without voice disorders.

Study	Control Group	Experimental Group	Speech Stimuli	Processing	RFF Offset 10 (ST)	RFF Onset 1 (ST)	Major Findings
	Participant Characteristics						
Arenas et al. [30]	Healthy children (*n*=14) 8 M, 6 F; *M*=41.9 mos range: [30, 57 mos]	Children who stutter (*n*=14) 9 M, 5 F; *M*=42.5 mos range: [30, 57 mos]	Continuous speech (Sentence)	Manual (CSL)	NR	NR	No group differences observed for RFF offset or onset cycles or their SDs. SDs varied by phonetic context (/apa/ vs. /itu/) for offset and onset.
Buckley et al. [2]	n/a	[a]AdLD (*n*=19) 9 M, 10 F; *M*=58.2 yrs, range: [37, 80 yrs]	Continuous speech (Sentence)	Manual (Praat)	*M*=−1.17 SD=1.49 range: [−3.51, 1.76]	*M*=2.55 SD=1.49 range: [−0.80, 4.60]	A combination of RFF offset and onset measures, including changes between RFF cycles, significantly predicted listener ratings of overall dysphonia.
Dahl and Stepp [31]	Healthy Adults; (*n*=20) 10 M, 10 F; *M*=20.2 yrs, SD=1.4 yrs; range: [18, 22 yrs]	n/a	Continuous speech (Stroop task)	Manual (Praat)	Congruent condition: *M*=−0.17 Incongruent condition: *M*=−0.42	Congruent condition: *M*=3.24 Incongruent condition: *M*=2.90	There was only a small effect of cognitive load on RFF offset cycles as a whole; no effect on onset cycles. Only 60% of subjects had the expected change.
Eadie and Stepp [4]	n/a	[a]AdLD (*n*=19) 9 M, 10 F; *M*=58.2 yrs, SD=10.8 yrs	Continuous speech (Sentence)	Manual (Praat)	Entire group: *M*=−1.2, SD=1.5 Subgroups: Mild effort: *M*=−0.82, SD=1.45; Moderate effort: *M*=−1.32, SD=1.33; Severe effort: *M*=−1.34, SD=2.20	Entire group: *M*=2.6, SD=1.5	RFF offset cycle 10 was not related to overall dysphonia severity or vocal effort, whereas onset cycle 1 was moderately related to both. Compared to a control group [13], RFF values were significantly lower for the AdLD group.
Fujiki et al. [32]	Healthy Adults; (*n*=16) 8 M, 8 F; *M*=22 yrs, range: [18, 28 yrs]	n/a	VCV	Automated (aRFF)	NR	NR	Vocal loading, with or without environmental humidity, had no impact on offset cycle 10 or onset cycle 1.
Goberman and Blomgren [5]	Aging Adults; (*n*=8) 5 M, 3 F; *M*=69.1 yrs	PD (*n*=9) 6M, 3 F; *M*=69.1 yrs, range: [57, 84 yrs]	Continuous speech (Rainbow passage)	Manual (CSL)	[b]Aging Adults: *M*=−1.1 [b]PD: *M*=−2.2	[b]Aging Adults: *M*=5.5 [b]PD: *M*=1.75	Control subjects showed significantly different offset and onset values when compared to the PD group, regardless of whether PD subjects were on or off medication. However, medication showed no effect within the PD group.
Heller Murray et al. [34]	Healthy Adults; (*n*=6) 3 M, 3 F; *M*=20.8 yrs,	Adults with high-voice use; (*n*=12) 12 F;	Continuous speech (Sentences)	Manual (Praat)	Healthy Adults: Pre: *M*=−1.0, SD=0.8 Post: *M*=−0.7, SD=1.1	Healthy Adults: Pre: *M*=2.9, SD=0.7 Post: *M*=2.9, SD=0.4	The experimental group showed significantly lower onset cycle 1 following 10-weeks of high voice use, but not

	SD=1.5 yrs	M=20.0 yrs, SD=1.5 yrs				High-voice use: Pre: $M=-0.9$; SD=1.2 Post: $M=-0.6$; SD=1.0	High-voice use: Pre: $M=3.2$, SD=1.6; Post: $M=2.5$, SD=1.6	changes in offset cycle 10. No changes were noted over the 10-week waiting period for control subjects. An ROC analysis of the change in onset cycle 1 (pre/post) showed 83% specificity and 58% sensitivity at -0.68 ST. Change in offset cycle 10 was deemed to be chance.
Heller Murray et al. [35]	Healthy children; (n=28) 15 M, 13 F; M=8.9 yrs	Children with vocal fold nodules; (n=28) 15 M, 13 F; M=9.0 yrs	Continuous speech (CAPE-V)	Manual (Praat)	NR	NR	No impact of group on RFF values; however, there was significantly less variability in offset cycle 10 for children with nodules. Offset cycle 10 was lower in males compared to females and onset cycle 1 was lower in younger children compared to older children.
Heller Murray et al. [6]	Healthy Adults; (n=49) 12 M, 37 F; M=39.2 yrs, SD=18.3 yrs	PVH (n=54) 7 M, 47 F; M=31.1 yrs, SD=14.7 yrs NPVH (n=35) 10 M, 25 F; M=41.7 yrs, SD=13.7 yrs	VCV	Automated (aRFF)	Healthy Adults: $M=-0.82$ PVH: $M=-1.76$ NPVH: $M=-1.35$	Healthy Adults: $M=1.68$ PVH: $M=1.18$ NPVH: $M=1.59$	Offset cycle 10 was significantly lower for PVH compared to NPVH and control subjects, as well as lower for NPVH compared to controls. Onset cycle 1 was significantly lower for PVH compared to NPVH and controls, however, there was no difference in onset 1 between NPVH and controls. ROC analysis showed low sensitivity for using offset cycle 10 and onset cycle 1 to distinguish between groups.
Kagan and Heaton [36]	Healthy Adults; (n=16) 5M, 11F; M=26 yrs, SD=3.7 yrs	n/a	VCV	Manual (Praat)	Pre-fatigue: $M=-1.10$, SD=0.53; Post-fatigue: $M=-1.31$, SD=0.56	Pre-fatigue: $M=2.52$, SD=0.82 Post-fatigue: $M=2.21$, SD=0.62	On average, offset cycle 10 and onset cycle 1 decreased after vocal loading. Participants who received low-level red-light therapy returned to baseline RFF levels faster that groups treated with infrared light, heat, or no treatment (control) following fatigue.
Lien et al. [38]	Healthy Adults; (n=16) 8M, 8F; M=21.5 yrs, SD=2.8 yrs	n/a	Continuous Speech (Sentences) and VCV)	Manual (Praat)	NR	NR	VCV utterances resulted in higher RFF offset cycle 10 compared to sentences, but lower onset cycle 1 compared to sentences. VCV utterances also showed lower variation (SD) compared to sentences. Phonemes /ʃ/, /s/, and /f/ have higher average values compared to other plosives/fricatives, but lower SD.
Lien et al. [7]	Healthy Adults; (n=12) 5M, 7F; M=22 yrs, SD= 2.7 yrs	n/a	Continuous speech (pa-trains, sentences, with modulation of effort/strain)	Manual (Praat)	NR	NR	Across subjects, RFF offset cycle 10 and onset cycle 1 showed poor relationships with an aerodynamic ratio (sound pressure level [dB SPL] to intraoral pressure) and listener perceptual ratings of vocal effort. When examined within-subject, relationships were moderate, but highly variable across subjects.

Study	Participants		Speech Task	Method	Onset Cycle 1	Offset Cycle 10	Findings
Lien and Stepp [41]	Healthy Adults; (*n*=12) 6M, 6F; range: [18, 28 yrs]	n/a	Continuous speech (sentences)	Manual (Praat)	*M*=−0.45	*M*=2.6	RFF averages as a whole as well as RFF offset cycle 10 from the accelerometer signals were significantly lower than those for an unprocessed microphone signal and the low-pass filtered microphone signal. Overall, SDs for the unprocessed microphone were lower than the other two signals.
Lien et al. [37]	Healthy Adults; (*n*=20) 2M, 18F; *M*=33 years, SD=17; range: [18, 87 yrs]	VH (21 PVH, 19 NPVH) (*n*=40) 5M, 35F; *M*=39 yrs, SD=17 yrs; range: [18, 75 yrs]	Continuous speech and VCV	Manual (Praat)	NR	NR	RFF calculated from a microphone was significantly higher than from an accelerometer, with *post hoc* testing showing that it was driven by cycles closest to the consonant. There was no impact of acquisition tool on RFF SDs.
McKenna et al. [43]	Healthy Adults; (*n*=18) 7M, 11F; *M*=33.72 yrs, SD=8.3 yrs; range: [23, 49 yrs]	n/a	VCV	Automated (aRFF)	Pre: *M*=−0.74, SD=0.81 Post: *M*=−1.01, SD=0.80	Pre: *M*=2.09, SD=1.29 Post: *M*=1.95, SD=0.64	A significantly lower RFF offset cycle 10 was found post-workday for occupational voice users. Interaction effects showed that men and subjects wearing N95 masks were driving the post-workday change. There were no significant findings for onset cycle 1.
McKenna et al. [42]	Healthy Adults; (*n*=21) 8M, 13F; *M*=32.9 yrs, SD=7.9 yrs; range: [23, 49 yrs]	n/a	VCV	Automated (aRFF)	Unmasked: *M*=−0.77, SD=0.90 Masked: *M*=−0.96, SD=1.05	Unmasked: *M*=2.10, SD=1.29 Masked: *M*=2.21, SD=1.35	RFF offset cycle 10 was significantly lower when participants wore face masks compared to unmasked speech. No significant effects for onset cycle 1.
McKenna and Stepp [8]	Healthy Adults; (*n*=26) 10M, 16F; *M*=20.9 yrs, SD=2.8 yrs; range: [18, 29 yrs]	n/a	VCV (with modulations of vocal effort)	Automated (aRFF)	Typical Speech: *M*=−0.51, SD=0.99 Mild Effort: *M*=−0.64, SD=1.08 Moderate Effort: *M*=−1.07, SD=0.91 Maximal Effort: *M*=−1.50, SD=1.29	Typical Speech: *M*=2.45, SD=0.81 Mild Effort: *M*=2.21, SD=0.60 Moderate Effort: *M*=2.23, SD=0.86 Maximal Effort: *M*=2.00, SD=0.77	RFF offset 10 significantly predicted listener perceptual ratings of vocal effort but did not predict speaker ratings. RFF onset 1 did not predict speaker or listener ratings of vocal effort.
McKenna et al. [1]	Healthy Adults; (*n*=12) 2M, 10F; *M*=22.7 yrs, SD=4.4 yrs; range: [18, 31 yrs]	n/a	VCV (with modulations of strain)	Automated (aRFF)	Typical Speech: *M*=−0.69, SD=0.8 range: [−1.34, 1.48]	Typical Speech: *M*=1.83, SD=0.98 range: [−0.32, 3.03]	Offset cycle 10 and onset cycle 1 were both significant predictors of kinematic estimates of laryngeal stiffness captured during modulations of strain. When examined by-subject, offset cycle 10 showed stronger relationships with kinematic stiffness ratios than onset cycle 1.
Park and Stepp [45]	Healthy Adults; (*n*=32) 16 M, 16 F	n/a	VCV	Automated (aRFF)	NR	NR	Offset cycle 10 and onset cycle 1 had moderate test/retest reliability over the recordings spaced over five days.

	M=22.5 yrs, SD=4.1 yrs; range: [18, 33 yrs] *Analysis on a subset of 28 subjects						Onset cycle 1 had the highest reliability ("good") when assessed in a soft voice, whereas offset cycle 10 had highest reliability ("good") during a loud voice.
Park and Stepp [10]	Healthy Adults; (*n*=15) 4M, 11F; *M*=21.9 yrs, SD=3.9 yrs; range: [19, 29 yrs]	n/a	VCV	Automated (aRFF)	NR	NR	Typical pitch and loudness, vowel /ə/, and equal syllable stress, reduced RFF variation (SD). Stress on the first syllable lowered offset cycle 10, whereas stress on the second syllable lowered onset cycle 1. Higher pitches decreased RFF values.
Park et.al [9]	Healthy Adults; (*n*=8) 4M, 4F; *M*=32.6 yrs, range: [18, 67 yrs]	n/a	VCV (with modulations of vocal effort)	Manual (STRAIGHT)	NR	NR	Synthetically lowering all 20 RFF values during typical phonation, increased the perception of strain. Synthetically raising all 20 RFF values during speech produced with vocal effort, reduced listener-perceived strain.
Park et al. [44]	Healthy Adults; (*n*=10) 5M, 5F; *M*=22.7 yrs; range: [19, 26 yrs]	Aging Adults; (*n*=10) 5M, 5F; *M*=62.1 yrs; range: [53, 76 yrs]	VCV	Manual (Praat)	NR	NR	There was no difference in RFF values at offset for /f/ vs. /t/, however, /f/ had larger abductory angles and longer durations to devoice. Onset cycle 1 was lower for /t/ compared to /f/, but /t/ had greater angles at onset and larger degrees of kinematic stiffness. RFF offset cycle 1 was moderately correlated with some adduction measures and kinematic stiffness ratios, but RFF offset 10 was only poorly correlated with abductory kinematics and stiffness ratios. There were no differences across young adults and aging adults.
Robb and Smith [46]	Healthy Adults; (*n*=10) 5M, 5F; *M*=21 yrs, 2 mos	Young Children; (*n*=10) 5M, 5F; *M*=4 yrs, 5 mos Older Children; (*n*=10) 5M, 5F; *M*=8 yrs, 5 mos	Continuous speech (Sentences)	Manual (CSL)	[d]Healthy Adults: *M*=−0.84	[d]Healthy Adults: *M*=2.8	Young adults showed higher onset cycle values (cycles 1 and 2) than young children. Older children showed significant lowering for offset cycles closest to the consonant, compared to the other two groups.
Roy et al. [47]	Healthy Adults; (*n*=20) 20 F; *M*=23.35 yrs, SD=7.01; range: [19, 52 yrs]	MTD (*n*=111) 111F; *M*=46.17 yrs, SD=13.69 yrs; range: [12, 79 yrs]	Continuous speech (Rainbow)	Manual (Praat)	NR	NR	MTD and controls were significantly different for onset cycles, but not for offset cycles. Following 1-hr of circumlaryngeal massage, the MTD group's onset cycles were similar to the control group (higher), indicating an improvement in onset cycle 1.

Study	Control group	Disordered group	Speech task	Method	Onset	Offset	Findings
							Onset cycle 1 was significantly correlated to dysphonia severity for the MTD group.
Smith and Robb [48]	Healthy Adults; (*n*=20) 10M, *M*=27 yrs; range: [22, 36 yrs] 10F; *M*=26 yrs; range: [22, 35 yrs]	n/a	VCV (Carrier phrase, "say __ again")	Manual (CSL)	NR	NR	Male participants had higher onset values compared to women. No differences for offset values noted. No differences between /p/ and /f/ phonemes noted.
Stepp et al. [13]	Healthy Adults; (*n*=15) 15 F; *M*=24.7 yrs, SD=2.8 yrs; range: [22, 33 yrs]	VH (*n*=82) 82 F; 1.Nodules: (*n*=30) *M*=32.8 yrs, SD=13.0 yrs; 2. MTD: (*n*=22) *M*=51.3 yrs, SD=18.5 yrs; 3. Polyps: (*n*=30) *M*=37.2 yrs, SD=11.5 yrs Pre/post sx, PVH (*n*=18) 18 F *M*=39.2 yrs, SD=11.2 yrs; range: [19, 61 yrs]	Continuous speech (Rainbow)	Manual (Praat)	[e]Healthy Adults: *M*=−0.33	[e]Healthy Adults: *M*=3.82	RFF values (all cycles) for control subjects were higher when compared to the VH group. Comparisons of RFF values pre/post-surgical removal of phonotraumatic lesions resulted in no changes.
Stepp et al. [24]	n/a	VH (*n*=16) 16 F (3 nodules; 13 MTD) *M*=32 yrs, SD=14 yrs; range: [18, 59 yrs]	Continuous speech (Rainbow)	Manual (Praat)	Pre-tx: *M*=−0.80, SE=0.26 Post-tx: *M*=−0.30, SE=0.21	Pre-tx: *M*=1.90, SE=0.18 Post-tx: *M*=2.71, SE=0.25	Post-therapy RFF cycles were significantly higher.
Stepp et al. [27]	Healthy Adults; (*n*=10) 10F; *M*=25 yrs, SD=11 yrs; range: [18, 53 yrs]	VH (*n*=30) 30F; *M*=31 yrs, SD=15 yrs; range: [18, 63 yrs]	Continuous speech (Rainbow)	Manual (Praat)	Healthy Adults: *M*=0.05, SD=1 range: [−1.78, 2.12] VH:	Healthy Adults: *M*=2.77, SD=0.87 range: [1.39, 3.89] VH:	Offset cycle 10 was significantly lower in VH compared to healthy controls, with no differences for onset cycle 1. Moderate negative correlations were reported between RFF values and listener-perceived overall dysphonia severity

					M=−1.75, SD=0.91 range: [−3.64, -0.40]	M=2.47, SD=1.03 range: [−0.36, 4.31]	and vocal effort for offset cycle 10, but weak correlations were found for onset cycle 1. ROC curve analysis identified a value of −0.56 ST with .97 sensitivity and .90 specificity for offset cycle 10 to distinguish between healthy and disordered speakers. ROC for onset cycle 1 was close to a chance detection rate.
Stepp [49]	Healthy Aging Adults; (n=32) 23M, 9F	PD (n=32) 23M, 9F; range: [43, 89 yrs]	Continuous speech (Rainbow and other sentences)	Manual (Praat)	Aging Adults: M=−1.6, SD=1.3 PD, off meds: M=−1.9, SD=1.2	Aging Adults: M=3.2, SD=1.2 PD, off meds: M=2.7, SD=1.2	RFF values were higher in females compared to males. RFF was lower in PD when on medication compared to controls, but no differences were found between PD on/off medication, or controls vs. PD off medication. Those with progressed PD had lower RFF values than those with new-onset.
Van Mersbergen and Lanza [50]	Healthy Adults; (n=26) 26F; M=23 yrs, SD=4.2 yrs; range: [20, 42 yrs]	n/a	VCV	Manual (Praat)	Positive condition: M=0.03, SD=0.10 Neutral condition: M=−0.01, SD=0.05 Aversive condition: M=0.29, SD=0.05	NR	Voicing during aversive emotional conditions resulted in higher offset cycle 10 values compared to neutral and positive conditions, however, there was no difference in neutral vs. positive. There was no difference in onset cycles, nor onset cycle 1 specifically.
Watson [12]	Healthy Young Adults; (n=10) 4M, 6F; M=24 yrs; range: [23, 27 yrs]	Aging Adults; (n=10) 4M, 6F; M=74 yrs; range: [68, 85 yrs]	Continuous speech (Sentence)	Manual (Voice+)	Young Adults: M=0.44 Aging Adults: M=−1.66	Young Adults: M=3.27 Aging Adults: M=2.80	Significant group differences for offset cycles 8, 9, and 10. No group differences for onset cycles.
Weerathunge et al. [53]	n/a	Voice Disorders (AdLD, PD, VH) (n=29) 15M, 14F M=51.8 yrs, SD=18 yrs; range: [19, 82 yrs]	VCV	Automated (aRFF-AP)	NR	NR	RFF could not be analyzed because there were too many missing values when audio signals were played through virtual telehealth platforms.

Note. AdLD: adductor laryngeal dystonia; PD = Parkinson's disease; VH= vocal hyperfunction; PVH= phonotraumatic vocal hyperfunction (e.g., nodules, polyp); NPVH = non-phonotraumatic vocal hyperfunction; MTD: muscle tension dysphonia; CSL = Computerized Speech Lab; VCV = vowel-consonant-vowel; M = male; F = female; M = mean; SD = standard deviation; SE=standard error; mos=months CAPE-V=Consensus Auditory-Perceptual Evaluation of Voice [65]; STRAIGHT = Speech Transformation and Representation using Adaptive Interpolation of weiGHTed spectrum. n/a = not applicable; NR=not reported; tx=therapy; sx = surgery; ROC=receiver operating characteristic. [a]The same subjects were analyzed in the studies by Buckley et al. [2] and Eadie and Stepp [4]. [b]=values are estimates reported in Stepp [49]; [c]Offset cycle 1 values were not reported in the study, but instead were gathered directly from authors. [d]=values calculated from information provided in the study, as raw values not reported. [e]=values reported in Stepp et al. [24].