

Title: Measuring vocal fatigue in sports coaches

Kathy Connaughton (BSc) NUI Galway

Irena Yanushevskaya (PhD) Trinity College Dublin and NUI Galway

Correspondence:

Dr Irena Yanushevskaya

Room 4091

Phonetics and Speech Laboratory

Centre for Language and Communication Studies

Arts Building

Trinity College Dublin

Dublin 2

phone: +353 1 896 1348

fax: +353 1 896 1560

yanushei@tcd.ie

Title: Measuring vocal fatigue in sports coaches

Abstract

Objective: This study explores the immediate impact of prolonged voice use by professional sports coaches.

Method: Speech samples including sustained phonation of [a] vowel and a short read passage were collected from two professional sports coaches. The audio recordings were made within an hour before and after a coaching session, over three sessions. The auditory analysis of the recordings was done using the GRBAS scale. The speech samples were subsequently analysed using Praat. The extracted acoustic measures included f_0 , jitter, shimmer, Harmonic-to-Noise ratio and Cepstral Peak Prominence. The data collected before and after coaching sessions was compared across speech types.

Main results: The results of auditory and acoustic analysis suggest a slight shift towards a tenser phonation post-coaching session, which is a likely consequence of laryngeal muscle adaptation to prolonged voice use. Differences in the levels of parameters measured were observed between sustained vowels and connected speech.

Conclusion: Selected acoustic measures and methodology can be useful to capture the subtle voice change post-coaching session. It is desirable, however, that more sophisticated and robust and at the same time intuitive and easy to use tools for voice assessment and monitoring be made available to clinicians and professional voice users.

Key words: vocal fatigue, sports coaches, acoustic analysis, f_0 , jitter, shimmer, HNR, Cepstral Peak Prominence.

Measuring vocal fatigue in sports coaches

Kathy Connaughton and Irena Yanushevskaya

Introduction

Voice disorders are becoming increasingly more prevalent (Colton, Casper, & Leonard, 2006), with occupational voice disorders attracting a lot of attention. A voice disorder can be defined as any abnormality with an individual's voice pitch, quality or loudness or any combination of these three. According to the American Speech-Language-Hearing Association, voice disorder 'is characterized by the abnormal production and/or absences of vocal quality, pitch, loudness, resonance, and/or duration, which is inappropriate for an individual's age and/or sex' (American Speech-Language-Hearing Association, 1993). No uniform standard for abnormal voice exists, just as there are no fixed criteria for normal voice. Voice disorders can be organic, functional or psychogenic. Organic voice disorders can be either structural or neurological. Functional voice disorders occur when the physical structure appears normal but the vocal mechanism is being misused. Psychogenic voice disorders occur when poor voice quality becomes a manifestation of unresolved psychological conflict (Gallena, 2007; Mathieson, 2001).

Professional or occupational voice users are those whose use of voice is of key importance in their job performance (Dejonckere, 2001; Sataloff, 2005; Verdolini & Ramig, 2001).

High levels of voice disorders have been reported amongst this population, with 20-80% of teachers and 46% of call centre workers reporting voice problems (Hazlett, Duffy, & Moorhead, 2011; Roy et al., 2004). In particular, they are susceptible to vocal fatigue.

Vocal fatigue is described as a negative vocal adaptation that occurs as a consequence of prolonged voice (Welham & Maclagan, 2003). Vocal fatigue is associated with undesirable

and unexpected changes in vocal quality such as laryngeal discomfort, hoarseness, breathy voice quality, reduced loudness, reduced vocal control and loss of voice (Solomon, 2008; Nanjundeswaran, Jacobson, Gartner-Schmidt, & Verdolini Abbott, 2015).

One group of professional voice users where objective evidence on vocal function is somewhat lacking is sports coaches. Sports coaches are at risk of vocal fatigue due to a number of factors. These include extended voice use at high vocal effort in order to be heard by players during training sessions and environmental factors such as weather conditions (Dejonckere, 2001). A number of questionnaire-based studies looked at the voice problems of group fitness instructors (Rumbach, 2013; Heidel & Torgerson, 1993; Newman & Kersner, 1998). A qualitative study by O'Neill and McMenamin (2014) is one of the first studies to explore the daily experiences of professional soccer manager's occupational voice use. However, only a handful of studies have been conducted using the objective acoustic measurements of the immediate impact of vocal strain in this population. A study by Wolf et al. (Wolfe, Long, Conner Youngblood, Williford, & Scharff Olson, 2002) reported that fitness instructors with self-reported voice problems used louder voice and their speech was characterised with higher jitter and lower Harmonic-to-Noise ratio (HNR) compared to the instructors without voice complaints, although no significant differences were found in the acoustic measures before and after the exercise session. A very recent study (Dallaston & Rumbach, in press) explored vocal performance of female group fitness instructors using an approach similar to the one presented here. Speech samples (sustained vowels, read speech and monologue speech) were collected from fitness instructors before and after a 60 minute exercise class and analysed in terms of f_0 , intensity, jitter, shimmer, HNR, maximum duration of sustained phonation, and pitch range. Additionally, self-ratings of voice quality were obtained before and after instruction. The study reports a significant increase in f_0 mean and minimum as

well as intensity post-instruction session in sustained vowels and read speech. Substantial cross-speaker variability was also reported. Another recent study of occupational vocal health in elite football coaches (Buckley, O'Halloran, & Oates, 2015) included measures of phonation time, f_0 mean and mode as well as vocal intensity during a coaching session using a portable phonation monitor. The study reported increased f_0 and vocal intensity during coaching sessions. Similar findings (e.g., increase in mean f_0 , decrease in jitter and shimmer) have been reported in teachers after a working day of voice use and in other studies investigating vocal loading, e.g., Laukkanen, Ilomaki, Leppanen, and Vilkman (2008). Overall, prolonged voice use in sports coaches seem to be associated with increase in f_0 (this has also been reported in many studies on vocal fatigue, e.g., Solomon (2008) and references therein), but the findings of perturbation measures are inconclusive.

The aim of the pilot study reported here is to establish (i) whether prolonged voice use (in a particular group of professional voice users – sports coaches) has any immediate impact on phonatory quality and (ii) whether acoustic analysis tools in a widely available free software package Praat (Boersma & Weenink, 2014) can measure this impact. The acoustic measures used in this study are selected from a range of analysis tools available in Praat, such as f_0 , perturbation measures (jitter, shimmer), and measures of periodicity of the voice (HNR, Cepstral Peak Prominence - CPP).

Acoustic measures of phonatory function are well established as tools in the area of clinical voice research as well as for objective evaluation of different treatment approaches and to monitor/evaluate the client's progress in the course of treatment (Hillenbrand, 2011; Baken & Orlikoff, 2000). Although a large amount of acoustic measures exist (Baken & Orlikoff, 2000; Kent & Ball, 2000; Maryn, Roy, De Bodt, Van Cauwenberge, & Corthals, 2009; Buder, 2000) in reality the most frequently used (in clinical context) are those readily available in software packages like Multi-Dimensional Voice Program or MultiSpeech

(Kay Elemetrics) or Voice Report in Praat (Boersma & Weenink, 2014). The present study uses the software and analysis techniques readily available to a practising speech and language therapist who often has no access to costly proprietary software (e.g., Multispeech). In particular, we use parameters such as f_0 mean and standard deviation, jitter and shimmer, HNR. These parameters are the ones most commonly used for the voice assessment and analysis in clinical research (Baken & Orlikoff, 2000) and are considered ‘the cornerstones of acoustic voice measurements’ (Kreiman & Gerratt, 2005, p.2201). In addition, the set of parameters used in this study includes CPP which is considered a more reliable measure of dysphonic voice than jitter and shimmer as they do not rely on pitch detection (Fraile & Godino-Llorente, 2014; Jannetts & Lowit, 2014; Maryn et al., 2009).

Acoustic analysis was preceded by an auditory analysis using GRBAS. Perceptual analysis of phonatory function using auditory scales is extensively used in clinic (Carding, Carlson, Epstein, Mathieson, & Shewell, 2000). GRBAS was selected in our study because this scale is the main auditory perceptual analysis tool used by speech and language therapists to assess the degree and quality of hoarseness (Omori, 2011). It is considered the minimum standard of perceptual analysis in the UK as adopted by the UK Royal College of Speech and Language Therapists (Carding et al., 2000). Auditory analysis is fundamentally important as auditory perceptual judgement is one of the key aspects of clinical decision making and provides a standard to compare objective measures to (Kent, 1996). Surveys demonstrate that clinicians rely heavily on auditory measures and that these methods are used frequently and are highly valued (Gerratt, Till, Rosenbek, Wertz, & Boysen, 1991).

Acoustic analysis of voice is usually performed on sustained isolated vowels. Although these are undoubtedly useful for the assessment of phonatory ability, they offer a rather poor representation of how voice functions in connected speech, in real-life situations of

conversational interactions. It has been noted that speakers may use compensatory strategies in sustained vowels that they may not be able to maintain in dynamically varying connected speech (Hillenbrand, 2011). The studies mentioned above found differences in the acoustic measurements of sustained vowels and connected speech. We therefore examine the difference between measures made on sustained vowels and connected speech.

The speech data was collected from two professional sports coaches before and after a one-hour coaching session and included sustained vowels and a short read passage. The number of participants in this study is admittedly very small, and no claims are made here about the generalisability of results. However, this study may act as a useful pilot and a basis for a larger study which would aim to further investigate the impact of prolonged voice use on this population using a much larger sample size.

Method

Participants

Before beginning this research, ethical approval was sought to ensure that this study conformed to the ethical practice guidelines of the Research Ethics Committee. Prospective participants were informed about the nature and purpose of the study and were required to sign a consent form which allowed their data to be utilised in this study.

The following criteria were used in the recruitment process: the participants were male sports coaches, aged between 20 and 40 years old who had at least two years of coaching experience. The participants were required to have no history of previous speech and language therapy to ensure that the results were representative of those who have not had

any voice therapy and have not been educated in voice care, as self-reported. The participants were required to be non-smokers and of good overall health.

As the goal of the study was to establish whether acoustic analysis tools readily available in Praat can capture the immediate impact on the voice that prolonged voice during a coaching session might have, we planned to record our participants within an hour before and after a coaching session, over three sessions, and the participants were asked to travel to the recording venue. The demands of the study were rather hard to meet for some of the prospective participants, and only two participants were recruited for the study. The participants were 24 and 36 years old. One of them had three years of experience while the other had nine. Both coaches worked part-time, on average four hours per week. They were soccer coaches who worked outdoors and did not use any kind of amplification device.

Although no formal self-assessment (e.g., with a questionnaire) was planned for the study before or after a coaching session, in an informal conversation with the researchers, the participants mentioned that the training session placed high demands on their voice as they were instructing a large group. One of the participants reported that his voice often felt strained after the coaching sessions.

Recording procedure

The recording of speech samples were made in a quiet (sound treated) room using a high quality Shure SM48 microphone held at about 15 cm from the speaker's mouth, directly onto a PC, at the sampling frequency of 44.1 kHz. The recording conditions were kept the same from session to session.

Speech material

The participants were recorded reading a short text, ‘The Rainbow Passage’ (Fairbanks, 1960) as well as sustaining the [a] vowel for about three seconds, and two repetition of each speech type were collected from each participant during each recording session. The recording sessions were held within one hour before and after a one-hour-long coaching session, over three coaching sessions.

The total data collected included: 2 speakers x 2 speech tasks (sustained [a], read passage) x 2 repetitions x 2 recording sessions (before/after) x 3 coaching sessions.

Data analysis

The recorded speech samples were initially assessed by a qualified speech and language therapist (over 20 years of clinical experience and a specialist in voice analysis) who was blinded to the aims of the study. The therapist carried out perceptual evaluation, using the GRBAS scale, on all the recorded samples. GRBAS was used for the auditory evaluation because as mentioned previously it is one of the most commonly used auditory analysis tool that has been shown to correlate with acoustic analysis measures (Bhuta, Patrick, & Garnett, 2004).

Subsequently, acoustic analysis of the data was carried out using Praat software (Boersma & Weenink, 2014). A number of acoustic parameters were extracted using Praat’s Voice Report, along with Cepstral Peak Prominence. The parameters are summarised below.

f_0 (Hz) : The rate of vocal fold oscillations, measured as the number of glottal cycles per second. We extracted f_0 mean and standard deviation as well as f_0 range (the difference between f_0 maximum and f_0 minimum values).

Jitter local (%): Frequency perturbation measure; defined in Praat as the average absolute difference between consecutive periods, divided by the average period; approximate threshold for pathology $> 1.040\%$.

Shimmer local (%): Amplitude perturbation measure; defined in Praat as the average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude; approximate threshold for pathology $> 3.810\%$.

HNR (dB): Harmonics-to-Noise ratio, a measure of the degree of the signal's periodicity; approximate threshold of pathology < 20 dB. In Praat, a HNR of zero means that there is equal energy in the harmonics and the noise.

CPPs (dB): Cepstral Peak Prominence (Hillenbrand et al., 1994; Hillenbrand & Houde, 1996) is a measure of the degree of periodicity of the voice signal. It is the difference in amplitude between the cepstral peak and the corresponding value on the regression line fitted to the cepstrum that is directly below the peak.

The measures of all the above parameters in sustained [a] vowels were made on steady-state portions of phonation (approximately three seconds in the middle of the vowel, excluding the onset and offset of phonation as recommended, e.g., in Baken and Orlikoff (2000). To get an indication of voice parameters in connected speech (read passage), we adopted an approach similar to the one described in Maryn and Weenink (2015) where acoustic measures of connected speech data were performed on automatically extracted and subsequently concatenated voiced speech segments. In our study, rather than selecting all voiced segments (which would include consonantal perturbations and might inflate

jitter and shimmer measures), we extracted 30 vowels from each recording of the read passage, concatenated them and conducted acoustic analysis on the concatenated signal. The vowels from the stressed (perceptually prominent) syllables were first manually segmented in Praat. To ensure unambiguous segmentation, we used vowels preceded and followed by obstruents, nasals or laterals (but not [ɹ], [w] or [j]). The extraction, concatenation and acoustic analysis were done automatically using a Praat script.

Results

Auditory analysis

The results of the auditory analysis are shown in Table 1. Although rather small, and in a number of cases not at all present, there are some directional changes in the perceptual evaluation of the speech samples: samples collected after coaching session are graded as showing decreased breathiness and asthenia, and increased strain, suggestive of a shift towards somewhat tenser phonation, as well as a slightly increased grade of hoarseness. The expectation therefore was that we would see patterns in the acoustic data suggestive of (slightly) tenser and hoarser phonation, such as, for example, increased jitter and shimmer values and decreased HNR and CPP values.

[Insert Table 1 about here]

Acoustic analysis

The results of the acoustic analysis are presented in Figure 1 and Table 2. Shown are the data collected before and after a coaching session, averaged over the three coaching sessions, for each speaker.

[Insert Figure 1 and Table 2 about here]

f_0 – There was a very slight increase of mean f_0 (2-3 Hz in sustained vowels, 3-6 Hz in connected speech) in the speech samples recorded post-coaching session compared to pre-coaching session samples, in both speakers in both speech types. As expected, higher mean f_0 and f_0 range were found in connected speech than in sustained vowels. There was no difference in f_0 range between pre- and post-coaching session recordings in sustained vowels. In connected speech, there was a slight increase in f_0 range post-coaching session (8 Hz and 4 Hz in Speaker 1 and Speaker 2 respectively).

Jitter – The jitter (local) measures showed a slight decrease (0.04%) in both sustained phonation and connected speech post-coaching session in both speakers. The jitter values in connected speech were higher than in sustained vowels and were above the 1.040% threshold of pathology for both speakers.

Shimmer – Shimmer (local) was also lower in both sustained vowels and connected speech in the samples collected post-coaching session in both speakers. Just like in jitter, there is a pronounced difference in the shimmer levels in sustained vowels and connected speech, with jitter being noticeably higher in the latter. For example, shimmer local was 2.10-2.26% post coaching session in sustained vowels, and 7.61-7.34% in connected speech. The level of shimmer (local) was above the 3.81% threshold of pathology in sustained [a] vowel post-coaching session (Speaker 1) and in connected speech (both speakers) in the samples collected pre- and post-coaching session.

HNR – The HNR measurements did not show any difference between the pre- and post-coaching session speech samples in both speakers. HNR in connected speech (12-15 dB) was found to be much lower than in sustained vowels (23-24 dB) and is indicative of an

elevated amount of noise in the speech signal (the values in connected speech are lower than the approximate threshold of pathology 20 dB).

CPP – There was a slight increase (1-2 dB) in CPP values in the speech samples collected post-coaching session in sustained vowels and connected speech (except in Speaker 2 connected speech where there was no difference in the pre- and post-coaching session samples). CPP values were lower in connected speech samples than in sustained vowels for both speakers.

Although the amount of data collected was small, an indication of overall differences in the speech samples collected before and after coaching sessions can be obtained by pooling the acoustic parameter data across speakers and speech tasks and comparing them using a paired samples t-test. The results of the test revealed that mean f_0 was significantly higher after the coaching session ($M = 115$, $SD = 8$) than before coaching session ($M = 111$, $SD = 6$), $t(11) = -2.51$, $p = 0.03$. Shimmer was significantly lower post-session ($M = 4.83$, $SD = 2.80$) than pre-session ($M = 6.39$, $SD = 2.76$), $t(11) = 2.59$, $p = 0.02$. The differences in jitter, HNR and CPP pre- and post-coaching session were not found statistically significant: jitter pre-session ($M = 0.75$, $SD = 0.61$), jitter post-session ($M = 0.72$, $SD = 0.61$), $t(11) = 0.37$, $p = 0.71$; HNR pre-session ($M = 18.6$, $SD = 5.4$), HNR post-session ($M = 19.6$, $SD = 5.4$), $t(11) = -1.57$, $p = 0.14$; CPP pre-session ($M = 27.0$, $SD = 3.5$), CPP post-session ($M = 27.9$, $SD = 4.0$), $t(11) = -2.19$, $p = 0.05$. Overall, the results of acoustic measurements showed rather minor changes in phonation post-coaching session compared to pre-coaching session.

Discussion

The results of both the auditory and the acoustic analysis suggest rather minor changes in phonation post-coaching sessions in the participants analysed here. GRBAS ratings conducted prior to the acoustic analysis suggests a shift towards somewhat tenser phonation after a prolonged voice use in both speakers (decreased asthenia and increased strain), although the ratings are within the mild range. This is in keeping with informal self-reports of the participants: an increased vocal strain after a coaching session. As the results of the auditory analysis showed, the changes in phonatory quality observed here are perceptually rather subtle, so it is not surprising that the acoustic measures did not show any striking differences. Contrary to expectations, the perturbation measures jitter and shimmer did not increase significantly but rather showed a tendency towards decrease. Increased mean f_0 , decreased jitter and shimmer and increased CPP and HNR in the speech samples collected post-coaching sessions are concordant with a shift towards a less breathy/more tense phonation mode than pre-coaching session. Our findings are similar to the ones reported by Dallaston and Rumbach (in press) who found significant increase in f_0 mean in group fitness instructors after a 60 minute exercise class. Similar findings (e.g., increase in mean f_0 , decrease in jitter and shimmer) have been reported in teachers after a working day of voice use and in other studies investigating vocal loading (e.g., Laukkanen et al., 2008 and references therein). Decreased perturbation values and increase in f_0 seem to be the result of laryngeal muscle adaptation to prolonged voice use. It is likely that the same is reflected in our findings here.

These findings appear to be counterintuitive. As discussed in Dallaston and Rumbach (in press), a common consequence of vocal loading is vocal fold oedema – the swelling of the vocal folds. An increased mass of the vocal folds would cause a decrease in f_0 and increase in the perturbation measures. Contrary to predictions, the amount of vocal loading/vocal

strain experienced by our participants did not seem to cause vocal deterioration. One possible explanation is the duration of vocal strain. It is not clear for what proportion of the one hour coaching session the participants used their voice, at what vocal effort and pitch. It would have been necessary to include measures of relative time speaking, as well as intensity and f_0 (e.g., percent phonation time and other measures used in Buckley et al., 2015) to estimate that. It is possible that one hour coaching session may not have been long enough to cause vocal trauma. Another possibility is that the gap between the coaching session and the recording session (although it was never more than one hour) was sufficient to allow recovery of the vocal function.

There was a noticeable difference between the levels of parameters measured in sustained vowels and connected speech. Measuring jitter and shimmer in connected speech is problematic due to the dynamic changes in the speech signal and consonantal perturbations (Baken & Orlikoff, 2000). In this study, we followed the approach suggested in Maryn and Weenink (2015) and took measurements from concatenated vowel segments. However, the increased level of perturbation measures and decreased HNR in connected speech could simply be an artefact of the concatenation procedure, and as a result it is not possible to directly compare the measurements for sustained vowels and connected speech. The overall tendencies in phonation shifts pre- and post-coaching were nevertheless similar in sustained vowels and connected speech.

The research questions were (i) whether prolonged voice use in sports coaches has any impact on the vocal quality and (ii) whether the available tools can register this impact. The answer to the first question appears to be yes, although the impact is rather subtle perceptually and in a rather unexpected direction: the phonation post-coaching session appears to be more efficient. The available acoustic parameters (f_0 , jitter, shimmer, HNR

and CPP) all reflected this shift in phonatory quality, but the differences were for the most part not significant statistically.

The reliability and clinical relevance of popular measures, such as jitter and shimmer have been questioned (Hillenbrand, 2011; Heman-Ackah et al., 2003; Kreiman & Gerratt, 2005). Both jitter and HNR depend on the accurate estimation of individual glottal pulses, which is problematic if the voice is even mildly dysphonic (Hillenbrand, 2011). A measure of periodicity of the voice signal - Cepstral Peak Prominence has been increasingly used in clinical voice research as a more robust and reliable measure of dysphonia (Hillenbrand, Cleveland, & Erickson, 1994; Heman-Ackah et al., 2003; Jannetts & Lowit, 2014; Maryn et al., 2009), however it did not show significant differences between the voice samples analysed here. It is possible that the acoustic measures used here are not well suited to register the relatively subtle difference in phonation.

Conclusions

This pilot study explored the immediate impact of prolonged voice use in sports coaches, and it is one of the first studies to acoustically analyse the voices of sports coaches in Ireland. Two professional sports coaches were recorded producing sustained vowels and connected speech (a short text) within an hour before and after a coaching session, over three sessions. The collected data was analysed using perceptual and acoustic measures.

The results suggest a mild shift towards tenser phonation post-coaching session, which most likely reflects increased laryngeal muscle activity as an adaptation after a prolonged voice use during one hour coaching session. Due to the complex design of the study (the recordings were done of two types of speech, pre/post coaching sessions, over three sessions), and the demands on the participants, only two participants were recruited for the

study. The number of participants is admittedly small, and a larger sample is necessary before any generalisations can be made. There are several ways in which this study can be extended. For example, the small sample size in this study precludes any formal statistical evaluation of the correlations between the perceptual and acoustic analyses. Future work could explore this further with a larger sample size. The participants acted as their own ‘baseline’ as the data pre-and post-coaching session was compared. It would be informative to include the data on self-evaluation of the participants’ voice, e.g., by using a questionnaire both before and after a coaching session. Including the information on the proportion of the coaching time spent vocalising as well as the data on the intensity and f_0 used by the sports coaches during a coaching session would also be necessary to estimate the actual vocal load. Different results for the [a] vowel and connected speech corroborate the view that the sustained vowels are not sufficient to represent vocal function and that connected speech (read or spontaneous) needs to be included in the analysis. The problem of reliability and validity of perturbation and noise measures in connected speech still remains.

Although the auditory and acoustic measures point at rather mild increase in the vocal strain post-coaching session here, this mild strain has a tendency to build over time and lead to more pronounced voice problems (e.g., Solomon, 2008). According to Verdolini and Ramig (2001), individuals using their voice more frequently than others and at higher demands are at higher risk of vocal injury. Monitoring and early detection of vocal strain in professional voice users would be beneficial.

As mentioned above, this study used a set of tools and acoustic measures readily available to a practising speech and language therapist. Selected acoustic measures and methodology can be useful to capture the subtle voice changes resulting from an increased vocal load.

However, they are not without limitations: their use requires a certain level of expertise in acoustic analysis and if done manually can be rather time consuming. Furthermore, they may not be well suited to the subtle type of differences we are dealing with here. As discussed in Solomon (2008), vocal fatigue may be experienced by the speaker, but the changes in phonation associated with it may not be audible or easily detected by instrumental measures. It is therefore desirable that more sophisticated and robust and at the same time intuitive and easy to use tools for voice assessment and monitoring be made available to both the public (professional voice users) and practising SLT clinicians.

Acknowledgements

The authors would like to thank the editor and two anonymous reviewers for their comments and suggestions.

References

- American Speech-Language-Hearing Association. (1993). Definitions of Communication Disorders and Variations. doi:<http://10.1044/policy.RP1993-00208>
- Baken, R. J., & Orlikoff, R. F. (2000). *Clinical Measurement of Speech and Voice* (2 ed.): Thomson Learning.
- Bhuta, T., Patrick, L., & Garnett, J. D. (2004). Perceptual evaluation of voice quality and its correlation with acoustic measurements. *Journal of Voice*, 18(3), 299-304.
- Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer (Version 5.3.85). Retrieved from <http://www.praat.org/>
- Buckley, K. L., O'Halloran, P. D., & Oates, J. M. (2015). Occupational vocal health of elite sports coaches: Anexploratory pilot study of football coaches. *Journal of Voice*, 29(4), 476-483. doi:<http://dx.doi.org/10.1016/j.jvoice.2014.09.017>

- Buder, E. H. (2000). Acoustic analysis of voice quality: a tabulation of algorithms 1902-1990. In R. D. Kent & M. J. Ball (Eds.), *Voice Quality Measurement* (pp. 119-244): Singular Publishing Group.
- Carding, P., Carlson, E., Epstein, R., Mathieson, L., & Shewell, C. (2000). Formal perceptual evaluation of voice quality in the United Kingdom. *Logopedics Phoniatics Vocology*, 25(3), 133-138.
doi:<http://dx.doi.org/10.1080/14015430050175860>
- Colton, R. H., Casper, J. K., & Leonard, R. (2006). *Understanding voice problems: a physiological perspective for diagnosis and treatment*. London: Lippincott Williams & Wilkins.
- Dallaston, K., & Rumbach, A. F. ([in press]). Vocal performance of group fitness instructors before and after instruction: changes in acoustic measures and self-ratings. *Journal of Voice*. doi:<http://dx.doi.org/10.1016/j.jvoice.2015.02.007>
- Dejonckere, P. H. (Ed.) (2001). *Occupational voice: care and cure*. The Hague, the Netherlands: Kugler Publications.
- Fairbanks, G. (1960). *Voice and articulation drillbook*. New York: Harper & Row.
- Fraile, R., & Godino-Llorente, J. I. (2014). Cepstral peak prominence: A comprehensive analysis. *Biomedical Signal Processing and Control*, 14, 42-54.
doi:<http://dx.doi.org/10.1016/j.bspc.2014.07.001>
- Gallena, S. K. (2007). *Voice and laryngeal disorders*. St Louis, Missouri: Mosby Elsevier.
- Gerratt, B. R., Till, J., Rosenbek, J. C., Wertz, R. T., & Boysen, A. E. (1991). Use and perceived value of perceptual and instrumental measures in dysarthria management. In C. A. Moore, K. M. Yorkston, & D. R. Beukelman (Eds.), *Dysarthria and Apraxia of Speech* (pp. 77-93). Baltimore: Brookes.
- Hazlett, D. E., Duffy, O. M., & Moorhead, S. A. (2011). Review of the impact of voice training on the vocal quality of professional voice users: implications for vocal

- health and recommendations for further research. *Journal of Voice*, 25(2), 181-191.
doi:<http://dx.doi.org/10.1016/j.jvoice.2009.08.005>
- Heidel, S. E., & Torgerson, J. K. (1993). Vocal problems among aerobic instructors and aerobic participants. *Journal of Communication Disorders*, 26(3), 179-191.
doi:[http://dx.doi.org/10.1016/0021-9924\(93\)90007-W](http://dx.doi.org/10.1016/0021-9924(93)90007-W)
- Heman-Ackah, Y. D., Heuer, R. J., Michael, D. D., Ostrowski, R., Horman, M., Baroody, M. M., . . . Sataloff, R. T. (2003). Cepstral peak prominence: a more reliable measure of dysphonia. *Annals of Otology, Rhinology and Laryngology*, 112(2), 324-333. doi:<http://dx.doi.org/10.1177/000348940311200406>
- Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech and Hearing Research*, 37(August), 769-778.
doi:<http://dx.doi.org/10.1044/jshr.3704.769>
- Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *Journal of Speech and Hearing Research*, 39, 311-321. doi: <http://dx.doi.org/10.1044/jshr.3902.311>
- Hillenbrand, J. M. (2011). Acoustic analysis of voice: a tutorial. *SIG 5 Perspectives on Speech Science and Orofacial Disorders*, 21(2), 31-43.
doi:<http://dx.doi.org/10.1044/ssod21.2.31>
- Jannetts, S., & Lowit, A. (2014). Cepstral analysis of hypokinetic and ataxic voices: correlations with perceptual and other acoustic measures. *Journal of Voice*, 28(6), 673-680. doi:<http://dx.doi.org/10.1016/j.jvoice.2014.01.013>
- Kent, R. D. (1996). Hearing and believing: some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology*, 5, 7-23. doi:<http://10.1044/1058-0360.0503.07>
- Kent, R. D., & Ball, M. J. (2000). *Voice Quality Measurement*: Singular Publishing Group.

- Kreiman, J., & Gerratt, B. R. (2005). Perception of aperiodicity in pathological voice. *Journal of the Acoustical Society of America*, 117(4 Pt 1), 2201-2211.
doi:http://10.1121/1.1858351
- Laukkanen, A. M., Ilomaki, I., Leppanen, K., & Vilkman, E. (2008). Acoustic measures and self-reports of vocal fatigue by female teachers. *Journal of Voice*, 22(3), 283-289. doi:http://dx.doi.org/10.1016/j.jvoice.2006.10.001
- Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: a meta analysis. *Journal of the Acoustical Society of America*, 126(5), 2619-2634. doi:http://dx.doi.org/10.1121/1.3224706
- Maryn, Y., & Weenink, D. (2015). Objective dysphonia measures in the program Praat: Smoothed Cepstral Peak Prominence and Acoustic Voice Quality Index. *Journal of Voice*, 29(1), 35-43. doi:http://dx.doi.org/10.1016/j.jvoice.2014.06.015
- Mathieson, L. (2001). *Green and Mathieson's The Voice and its Disorders* (6 ed.). London and Philadelphia: Whurr Publishers.
- Nanjundeswaran, C., Jacobson, B. H., Gartner-Schmidt, J., & Verdolini Abbott, K. (2015). Vocal Fatigue Index (VFI): development and validation. *Journal of Voice*, 29(4), 433-440. doi:http://dx.doi.org/10.1016/j.jvoice.2014.09.012
- Newman, C., & Kersner, M. (1998). Voice problems of aerobics instructors: Implications for preventative training. *Logopedics Phoniatrics Vocology*, 23(4), 177-180.
doi:http://dx.doi.org/10.1080/140154398434086
- O'Neill, J., & McMenamin, R. (2014). Voice use in professional soccer management. *Logopedics, Phoniatrics, Vocology*, 39(4), 169-178.
doi:http://dx.doi.org/10.3109/14015439.2013.825008
- Omori, K. (2011). Diagnosis of voice disorders. *Japan Medical Association Journal*, 54(4), 248-253.

- Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., & Smith, E. M. (2004). Prevalence of voice disorders in teachers and the general population. *Journal of Speech, Language and Hearing Research*, 47(2), 281-293.
doi:[http://dx.doi.org/10.1044/1092-4388\(2004/023\)](http://dx.doi.org/10.1044/1092-4388(2004/023))
- Rumbach, A. F. (2013). Voice problems of group fitness instructors: diagnosis, treatment, perceives and experienced attitudes and expectations of the industry. *Journal of Voice*, 27(6), e1-9. doi:<http://dx.doi.org/10.1016/j.jvoice.2013.03.012>
- Sataloff, R. T. (2005). *Professional voice: the science and art of clinical care* (3 ed.). San Diego, CA: Plural Publishing, Inc.
- Solomon, N. P. (2008). Vocal fatigue and its relation to vocal hyperfunction. *International Journal of Speech-Language Pathology*, 10(4), 254-266.
doi:<http://dx.doi.org/10.1080/14417040701730990>
- Verdolini, K., & Ramig, L. O. (2001). Review: occupational risks for voice problems. *Logopedics, Phoniatrics, Vocology*, 26(1), 37-46.
doi:<http://dx.doi.org/10.1080/140154301300109125>
- Welham, N. V., & Maclagan, M. A. (2003). Vocal fatigue: current knowledge and future directions. *Journal of Voice*, 17(1), 21-30. doi:[http://dx.doi.org/10.1016/S0892-1997\(03\)00033-X](http://dx.doi.org/10.1016/S0892-1997(03)00033-X)
- Wolfe, V., Long, J., Conner Youngblood, H., Williford, H., & Scharff Olson, M. (2002). Vocal parameters of aerobic instructors with and without voice problems. *Journal of Voice*, 16(1), 52-60. doi:[http://dx.doi.org/10.1016/S0892-1997\(02\)00072-3](http://dx.doi.org/10.1016/S0892-1997(02)00072-3)

Figure 1. Acoustic measures in sustained [a] vowels and read passage. Shown are mean parameter values and SD (error bars) over three sessions. White bars = acoustic measurements of the data

collected before a coaching session, grey bars = acoustic measurements of the data collected after a coaching session. Black dashed lines indicate threshold of pathology (where available).

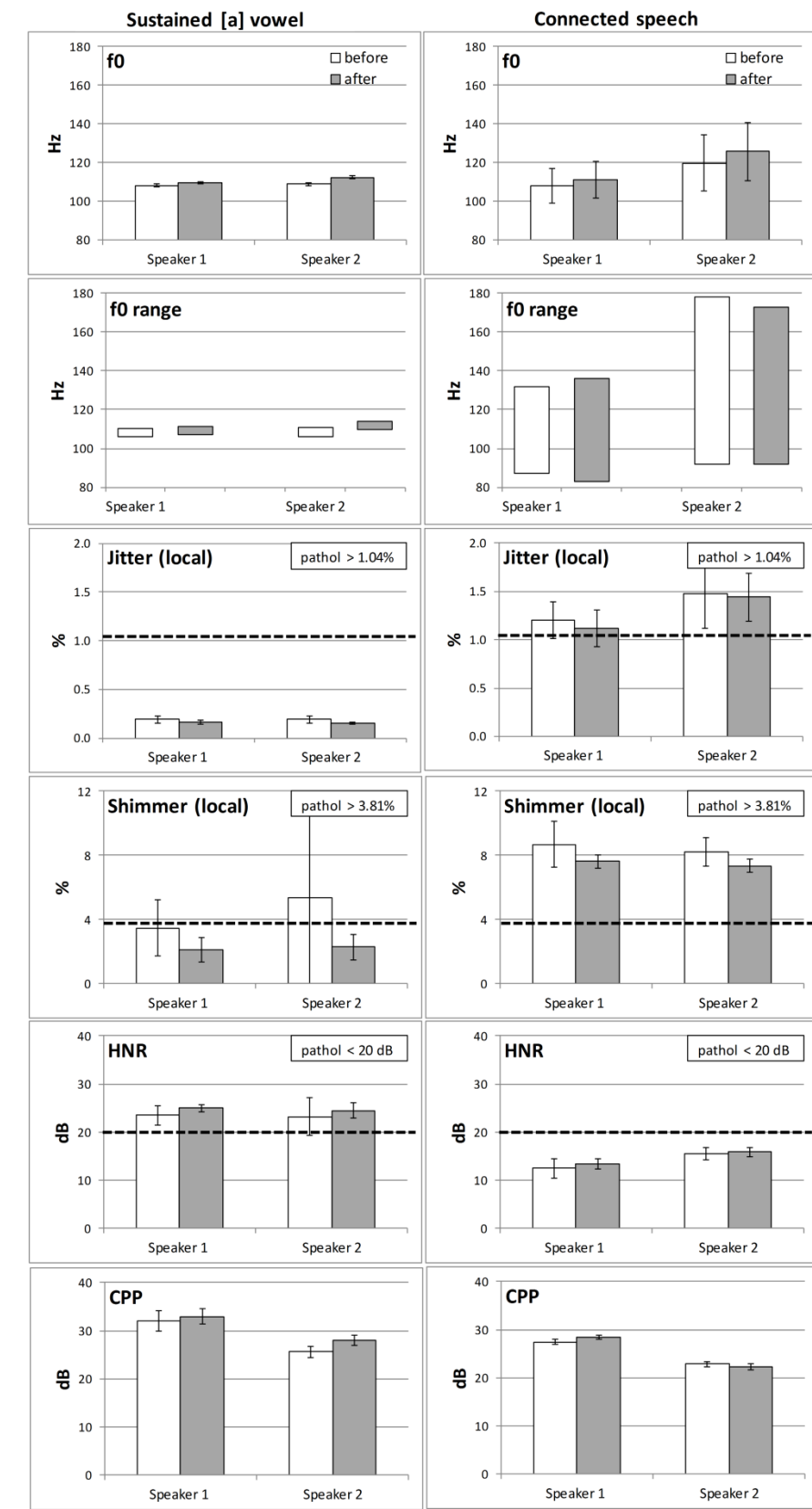


Table 1. Results of the auditory analysis using GRBAS. G=grade [of hoarseness], R=roughness, B=breathiness, A=asthenia, S=strain; 0=normal, 1=slight degree, 2=medium degree, 3=high

degree. 'Pre' = scores of the data collected before a coaching session; 'post' = scores of the data collected after a coaching session.

	Sustained [a] vowel											
	Speaker 1						Speaker 2					
	Session 1		Session 2		Session 3		Session 1		Session 2		Session 3	
	pre	post	pre	post	pre	post	pre	post	pre	Post	pre	post
G	1	1	1	1	1	2	0	1	0	0	0	1
R	0	0	0	0	0	0	0	0	0	0	0	0
B	1	0	1	0	1	1	0	0	0	0	0	0
A	0	0	1	0	0	0	0	1	0	0	0	1
S	0	1	0	1	0	1	0	0	0	0	0	0

	Connected speech											
	Speaker 1						Speaker 2					
	Session 1		Session 2		Session 3		Session 1		Session 2		Session 3	
	pre	post	pre	post	pre	post	pre	post	pre	Post	pre	post
G	1	2	1	1	0	0	1	1	1	1	0	1
R	0	0	0	0	0	0	0	0	0	0	0	0
B	1	2	1	0	1	2	1	1	1	0	0	0
A	0	0	1	0	0	0	0	0	1	0	0	0
S	0	1	0	1	0	0	0	1	0	0	0	1

Table 2. Acoustic measurements taken before and after coaching sessions. Shown are the values averaged across the three sessions and SD (in brackets).

Acoustic parameter	Sustained [a] vowel				Connected speech			
	Speaker 1		Speaker 2		Speaker 1		Speaker 2	
	pre	post	pre	post	pre	post	pre	post
f_0 (Hz)	108 (1)	110 (4)	109 (4)	112 (3)	108 (9)	111 (9)	119 (14)	125 (15)
f_0 range min-max (Hz)	106- 110	107- 111	106- 111	110- 114	87-133	83-137	92-178	92-174
Jitter local (%)	0.19 (0.04)	0.16 (0.02)	0.19 (0.04)	0.15 (0.01)	1.20 (0.19)	1.12 (0.19)	1.48 (0.36)	1.44 (0.25)
Shimmer local (%)	3.46 (1.73)	2.10 (0.75)	5.33 (5.45)	2.26 (0.79)	8.67 (1.44)	7.61 (0.43)	8.19 (0.90)	7.34 (0.39)
HNR (dB)	23.4 (1.9)	24.9 (0.7)	23.1 (3.9)	24.4 (1.6)	12.4 (2.0)	13.1 (1.0)	15.4 (1.2)	15.8 (0.9)
CPP (dB)	31.9 (2.1)	32.9 (1.5)	25.6 (1.1)	28.1 (1.1)	27.4 (0.5)	28.4 (0.4)	22.8 (0.5)	22.3 (0.6)