A Study of Subharmonics in Connected Speech Material

[*]Elena Kramer, [†]Roland Linder, and [*]Rainer Schönweiler, *[*†]Lübeck, Germany*

Summary: Subharmonics are often observed in running speech spectrograms that are difficult to quantify. This study investigates the relationship between rough voice quality and the presence—and amount—of subharmonics in connected speech material in a group of 35 male and 35 female speakers with voice pathology. Spectrum analysis was undertaken in 145 pathologic voices, of which 77 had subharmonics in connected speech. Only 34% of 70 subjects under investigation developed subharmonics in both connected speech and sustained phonations. All voices were judged as perceptually rough. The results of this study indicate that male and female voices with subharmonics do not differ in examined acoustic characteristics except for the modal fundamental frequency (F_0). A small but significant difference in roughness scores between genders seems to be caused by the confounding perception of factors not related to acoustic measurements of F_0 and number of subharmonics. The degree of roughness has a significant relationship with the irregularity index, percentage of low F_0 estimates, and, to a lesser extent, power of subharmonics.

Key Words: Subharmonics–Roughness.

INTRODUCTION

Subharmonics are whole-number divisors of the fundamental frequency (F_0) that can be seen as additional traces between the harmonics in the broadband spectrum and correspond to smaller distinct peaks located between two consecutive harmonic peaks in the power spectrum. Subharmonics present a periodically occurring "irregularity" (as opposed to randomly occurring irregularity measured by jitter and shimmer) that rarely persists throughout the duration of the vowel.[1–3] Up to three subharmonics can be distinguished between two consecutive harmonic traces.[2] Subharmonics, which have often been observed in connection with asymmetric vocal fold vibrations because of differences in mechanical properties of the vocal folds,[4] have been implicated as acoustic property of pathologic voices. Clinical studies reporting on subharmonics included cases with polyps, cysts, polypoid degeneration, plicae ventricularis, atrophy, and paralysis.

Considering a wide variety of pathologies that cause asymmetric vocal fold vibration or desynchronization of the vibratory modes of a single vocal fold, subharmonics are expected to be common in disordered voices. However, information on the occurrence of subharmonics in sustained phonations of voice-disordered population available from larger clinical studies do not substantiate this expectation. According to Núñez Batalla et al,[5] subharmonics were estimated in 31% (36 from 115) of examined pathologic voices. The presence of subharmonics in dysphonic vowels was reported to be significantly lower in the study by Omori et al.[6] Only 20 patients in 389 (5.1%) were found to have subharmonics. In the study by Behrman et al,[2] 9.4% of subjects (19/202) had subharmonics. Moreover, there is evidence that subharmonics can evolve in the absence of voice pathology. Wong et al[7] showed that subharmonics could be generated in vocal folds of normal stiffness and mass without asymmetry with decreased stress in the longitudinal string tension. Svec et al[8] described a subharmonic vibratory pattern in a normal larynx. Haben et al[3] observed mucosal wave asymmetry and subharmonics in subjects without voice pathology.

Prior clinical research has demonstrated perceptual salience of subharmonics in dysphonic vowel production. Dejonckere and Lebacq[1] and Wong et al[7] related subharmonics to the perception of diplophonia (two simultaneously perceived pitches). In Cavalli and Hirson,[9] 90% of vowels perceived as diplophonic had subharmonics. Klatt and Klatt[10] suggested that subharmonics evoke a sensation of creak characterized by perception of a combination of low pitch and roughness. The results of recent investigations by Núñez Batalla et al[5] and Omori et al[6] agree in that the presence of subharmonics in sustained phonations was found to signal roughness, whereas traditional measures of roughness such as jitter and shimmer were within normal range. It has been shown in experiments with synthetic signals that subharmonics (F_0/2) interfere with perceived pitch and contribute to perceived roughness.[11,12]

Roughness is a prominent feature in many pathologic conditions affecting the regularity of vocal fold vibrations irrespective of whether subharmonics are involved. In the study by Askenfelt and Hammarberg,[13] rough vocal quality is assumed to be perceived as a low-pitched noise caused by irregular vocal fold vibrations. Consistent with findings in studies investigating roughness and perceived pitch, there seems to be an interaction between perceived pitch and roughness. Emanuel and Smith[14] and Newman and Emanuel[15] have shown that perceived vowel roughness decreased as the pitch level was raised. In both studies, the researchers pointed to the possibility that perceived roughness may depend on the relative pitch within one's personal range than on absolute pitch. Wolfe and Ratusnik[16] found that vowels rated as moderately and severely dysphonic

Accepted for publication August 7, 2012.
From the *Department of Phoniatrics and Pediatric Audiology, University of Lübeck, Lübeck, Germany; and the †Institute of Medical Informatics, University of Lübeck, Lübeck, Germany.
Address correspondence and reprint requests to Elena Kramer, Department of Phoniatrics and Pediatric Audiology, University of Lübeck, Ratzeburger Allee 160, D-23538 Lübeck, Germany. E-mail: kramer.elena@googlemail.com
Journal of Voice, Vol. 27, No. 1, pp. 29-38
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2012.08.005

received significantly lower pitch match values (Spearman rho = −0.64) than the less dysphonic and normal vowels. Another finding of this study relates perceived pitch to spectrographic noise classification in dysphonic vowels ($r_s = -0.57$). According to this finding, pitch perception is influenced by spectral voice characteristics. Moreover, the results of Hanson and Emanuel[17] suggested that vowel spectral noise measures provide a reliable acoustic index of vocal roughness. Yet, the question of which acoustic parameters may serve to describe the degree of roughness severity due to subharmonics still remains to be answered. Among acoustic properties of subharmonics that have been brought in relation with the degree of perceived roughness are the frequency and magnitude of subharmonics. In the report by Omori et al,[6] the frequency of subharmonics was found to be inversely related to the degree of perceived roughness. Because reported subharmonic frequencies measured exactly one half of the F_0 in all but one study subject, one might assume the existence of a link between F_0 of phonation and perceived roughness. This point is supported by research evidence suggesting that low F_0 in general is associated with perceived roughness. Wendahl[18] showed that signals with a low F_0 tended to be perceived as rougher than those with a high F_0. Verdonck-de Leeuw and Mahieu[19] found that the habit of smoking, which has often been reported to be one of the major causes of voice pathology, has a lowering effect on F_0 and is accompanied by an increase in roughness. These findings were supported in the study by Bergan and Titze (2000),[11] who also found in pitch-matching task that pitch perception in signals with subharmonics is continuous: although subjects preferred the subharmonic pitch as the true pitch, they sometimes chose intermediate pitches between the harmonic and subharmonic pitch. Lower F_0 appeared to produce earlier identification of subharmonic pitch. From this evidence, voices of female subjects with subharmonics, who typically speak at higher frequencies, might have been expected to be rated as less rough. However, to our knowledge, no gender-specific differences have been reported with regard to perceived level of roughness, whereas a significant gender effect has been documented for acoustic measures of roughness such as jitter and shimmer.[20–22]

In experiments with synthetic signals, Omori et al[6] have shown that perceived roughness increased with increasing power of subharmonics. Sun and Xu[12] proposed an algorithm to quantify the amplitude ratio between harmonics and subharmonics. They found that when the ratio was smaller than 0.2, subharmonics had no effect on pitch perception. As the ratio exceeded a value of 0.4, the pitch corresponded mostly to the lowest subharmonic frequency. Similarly, obtaining correct F_0 values in signals with subharmonics is a challenge because when there are subharmonics, the subharmonic peak might be wrongly selected as a candidate for F_0. One of the most common methods to reduce the occurrence of subharmonic errors is using a gender-specific lower limit of the F_0 tracker. A more important question is, however, whether subharmonic errors might be exploited to extract voice parameters to indicate the perceptual quality of voice. In particular, there may be two other factors affecting perceived roughness in signals with subharmonics: the frequency with which subjects switch from one mode of vibration to another (with and without subharmonics) or the amount of signal contaminated with subharmonics. Obviously, this information can be to a certain extent elicited from F_0 contours and F_0 histograms. An interesting observation was made by Kotby et al[23] when inspecting F_0 profiles of sustained phonations in dysphonic and normal populations. In voice-disordered subjects, they frequently found bimodal and multimodal F_0 distributions that were attributed to the presence of subharmonics and low-frequency modulations. No attempt was made to describe these F_0 distributions in quantitative terms.

Most of what we know about subharmonics stems from research performed on stable signals. From these limited data, no general conclusions about how these findings apply to connected speech are warranted. It is conceivable that acoustic and perceptual data obtained from sustained vowels and connected speech fragments may differ quantitatively and qualitatively. Although the type of voice segment (vowel vs sentence) presented for evaluation has reportedly no effect on reliability and agreement between the raters,[24,25] the magnitude of the ratings seems to depend on the choice of the voice segment in some studies but not in the others. Askenfelt and Hammarberg (1981)[13] argued that vowels are not representative of voice function status. Only in the case of severe pathology was vocal function found to be consistent between vowels and sentences. Despite the strong correlation between vowels and sentences in perceptual quality ratings (r ranging from 0.72 to 0.89), Hanson and Emanuel[17] found that dysphonic patients occasionally produce vowels that are less severely disturbed than sentences. Conversely, Wolfe et al[26] reported that vowels from normal subjects were assigned greater severity ratings than sentences. They measured a relatively high correlation coefficient of 0.78 on dysphonic severity between vowels and sentences. Vowel ratings accounted for 61% of the variance in the prediction of sentence severity. There was no significant difference found in the ratings of complete vowels (including onsets and offsets) and connected speech in Revis et al.[25] Similarly, no significant difference in ratings between oral reading and the sustained vowel was found in the study by Zraick et al.[27] Up to date, the question remains unresolved as to what extent acoustic measures made on vowels can be expected to predict the perceptual quality of connected speech. Although most researchers agree that connected speech is more natural and measures obtained from connected speech are more likely to be generalizable to a patient's everyday speech, relatively little research has been undertaken to understand the relationship between instrumental measures obtained from connected speech material and voice pathology in general or perceptual voice quality.[28–33] Aspects of speech such as articulatory changes, high level of instability and lack of consistency in voice function, the ability to change the laryngeal tone and intensity according to the requirements of speech, voice breaks, and frequent onset and offset of voicing do not manifest themselves in an isolated vowel. Despite that speaking fundamental frequency (SFF) is one of the most reported acoustic voice parameters in voice research, vocal fold frequency measurement in