Mobile Phone Applications Voice Tools and Voice Pitch Analyzer Validated With LingWAVES to Measure Voice Frequency

[*,†]Lidewij S. Boogers, [*,†]Britney S.J. Chen, [†,‡]Marieke J. Coerts, [†,‡]Rico N.P.M. Rinkel, and [†,§]Sabine E. Hannema,
*†‡§*Amsterdam, The Netherlands*

Summary: Objectives. Voice frequency can be measured to assess the voice change in transgender men and women during treatment. There are many applications that can analyze voice frequency. This validation study aimed to compare the ability to measure voice frequency of the mobile phone applications "Voice Tools" and "Voice Pitch Analyzer" with the registration program LingWAVES as the gold standard.

Study Design. Prospective validation study.

Methods. A total of 45 participants of whom 20 transgender individuals were included. They were asked to read "The North Wind and the Sun" twice. The first measurement compared voice frequency measured by Voice Tools with LingWAVES while the second measurement compared Voice Pitch Analyzer with LingWAVES. The two applications that were being compared simultaneously measured the voice frequency. Pearson's regression correlations were performed to test for correlation between the mobile phone applications and LingWAVES.

Results. Significant correlations were demonstrated between the measurements of Voice Tools and LingWAVES ($P < 0.001$) and between Voice Pitch Analyzer and LingWAVES ($P < 0.001$). Voice Tools overestimated voice frequency with a median deviation of 2Hz (range -4 to 20). The overestimation was more pronounced in the high ranges. Voice Pitch Analyzer showed underestimation of voice frequency in high ranges. Median deviation was -2Hz (range -16 to 14).

Conclusions. This validation study shows that voice frequency can be reliably measured with the mobile phone applications "Voice Tools" and "Voice Pitch Analyzer". Combined with the ease of use, these applications can be used to measure voice frequency in clinical practice and for research purposes.

Key Words: Voice—Voice frequency—Transgender—Smartphone application.

INTRODUCTION

The voice is essential for vocal communication and therefore also important for social functions. Furthermore, voice characteristics are part of one's identity. The voice frequency is the main voice characteristic and often used to identify gender, since the male voice has a lower frequency than the female voice.[1,2]

Individuals with gender dysphoria have a gender identity that differs from the sex they were assigned at birth. Their voice frequency range may fit with the sex assigned at birth instead of their identified gender. In both trans men (female sex assignment, male gender identity) and transwomen (male sex assignment, female gender identity) this might lead to voice related gender dysphoria.[3] Gender affirming hormone treatment (GAHT) may be initiated to induce physical changes that match the gender identity. Feminizing treatment consists of estradiol and masculinizing treatment consists of testosterone.[4]

One of the effects of testosterone treatment is the lowering of the voice frequency.[5,6] Previous research has shown that one year of testosterone treatment in the majority of trans men leads to a voice that does not differ from the voice of men from the general population.[7] However, the treatment of transwomen for voice frequency is more complex. For them, treatment options are voice therapy or phonosurgery.[6,8,9]

During treatment of voice related gender dysphoria, voice frequency can be measured by voice clinicians to monitor if treatment goals are achieved. This diagnostic tool is also widely used in other people with voice concerns such as people with muscle tension dysphonia and vocal nodules.[10,11] The measurement of voice frequency is usually performed using professional voice measurement equipment which is only available in hospitals or specialized clinics. However, over the past years, an increasing number of mobile phone applications has been developed to measure voice characteristics.[12-14] The mobility, accessibility and ease of use of these applications in combination with the low costs, can make them a good alternative for research or clinical purposes when the use of professional equipment is not possible.[13,14] However, the use of these apps for clinical purposes can only be implemented if they are shown to produce reliable measurements. Incorrect acoustic measurements could lead

Accepted for publication October 17, 2022.

From the *Department of Endocrinology, Amsterdam University Medical Center, Amsterdam, North Holland, The Netherlands; †Center of Expertise on Gender Dysphoria, Amsterdam University Medical Center, Amsterdam, North Holland, The Netherlands; ‡Department of Otolaryngology, Amsterdam University Medical Center, Amsterdam, North Holland, The Netherlands; and the §Department of Pediatric Endocrinology, Amsterdam University Medical Center, Amsterdam, North Holland, The Netherlands.

Address correspondence and reprint requests to Sabine E. Hannema, PO Box 7057, 1007 MB Amsterdam, Department of Pediatrics, Section Endocrinology, Amsterdam University Medical Centers, Amsterdam, The Netherlands. E-mail: s.e.hannema@amsterdamumc.nl

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

https://doi.org/10.1016/j.jvoice.2022.10.013

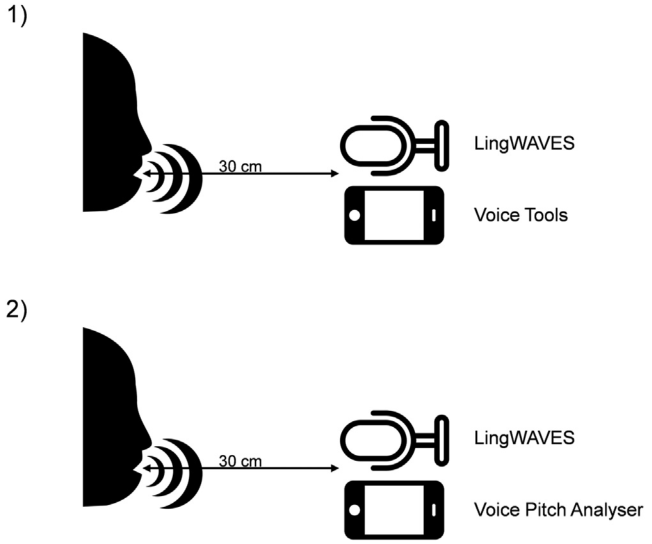

FIGURE 1. Set-up of the two measurements to validate (1) Voice Tools and (2) Voice Pitch Analyzer.

to incorrect diagnoses and/or unnecessary treatments. Several applications have been tested on their validity, which showed promising results.[12-15] Unfortunately, at this point, no free of charge applications have been validated.

Therefore, the aim of this study was to validate the ability of the free of charge mobile phone applications "Voice Tools" and "Voice Pitch Analyzer" to determine voice frequency with the registration program LingWAVES. Furthermore, differences in voice frequency between participant groups, ie transgender and cisgender males and females are reported to illustrate the range of voice frequency in transgender people.

METHODS

Participants

The study was carried out at the outpatient gender clinic from the hospital Amsterdam UMC location VUmc in May and June 2021. People with an appointment were invited to participate in the study as well as the people who accompanied them. All participants were required to speak Dutch and participants with voice disorders were excluded. Participants who visited the outpatient gender clinic are referred to as trans women or trans men. People who accompanied them and participated in the study are referred to as control men and control women.

Study protocol

The two free-of-charge applications that measured voice frequency that were available at the time of the study, Voice Tools and Voice Pitch Analyzer, were investigated. The participants were asked to read the text "De Noordenwind en de Zon" (*The North Wind and the Sun*) in a standing position to obtain voice samples (Appendix I). During the first measurement, voice characteristics were measured with LingWAVES and Voice Tools simultaneously. During the second measurement, voice characteristics were measured with LingWAVES and Voice Pitch Analyzer simultaneously. All the measurements took place in a quiet room in the hospital. The mobile phone applications were downloaded on an iPhone Xr. Both the internal microphone of the iPhone Xr and the microphone connected with LingWAVES were placed at 30 centimeters from the participants' mouth (Figure 1).

LingWAVES

The software program *LingWAVES* (Software Version 3.2, *Wevosys*) is used in speech therapy for voice measurements in our clinic. Since this program is broadly used by voice professionals this was used as the gold standard.[16,17] Therefore, voice measurements were performed twice with LingWAVES. A stand-based microphone was used. LingWAVES measured the mean, standard deviation, lowest, highest and range of both frequency and volume.

Voice tools

The application Voice Tools measured the average, median, highest (95^{th} percentile), and lowest frequency (95^{th} percentile) in Hz. Besides, it also measured the average, median, highest (95^{th} percentile), lowest (95^{th} percentile) of the volume. Environmental volume was also measured.

Voice pitch analyzer

The application Voice Pitch Analyzer measured the lowest frequency, median frequency and highest frequency in Hz.

Statistical analysis

Statistical analyses were performed using *Stata* version 15.1. Normally distributed data were summarized as mean with standard deviation (SD) while not normally distributed data were summarized as median with interquartile range (IQR). Pearson correlation coefficient was used to assess correlation between the two LingWAVES voice frequency measurements as well as between the mobile phone application measurements and LingWAVES measurements. This was performed for lowest frequency, highest frequency, average frequency, lowest volume, highest volume and average volume.

Linear regression analyses were performed to assess differences in average frequency between the different study groups.

Ethics

The study was approved by the local medical ethical committee. Informed consent was obtained from all participants included in the study.

TABLE 1.
Characteristics of the Study Population.

	Control Men	Control Women	Transwomen	Transmen	Total
Subjects, n (%)	7 (16)	18 (40)	11 (24)	9 (20)	45
Age, years (IQR)	51 (26 to 59)	27 (23 to 47)	25 (17 to 57)	19 (18 to 26)	26 (19 to 50)
GAHT, n (%)	-	-	11 (100)	7 (78)	18 (40)

Abbreviations: GAHT,gender affirming hormone treatment; IQR, inter quartile range

RESULTS

A total of 45 participants (20 transgender individuals, 25 controls) were enrolled in this study. Table 1 shows an overview of the characteristics of the study population. The median age at participation was 26 years (range 12 to 71). Except for two trans men, all the transgender participants received GAHT.

Validation

Measurements of average, lowest and highest frequency from both Voice Tools as Voice Pitch Analyzer were very closely correlated with the measurements performed by LingWAVES. These correlations were similar to those between the first and second measurement using LingWAVES. An overview of the correlation coefficients is shown in Table 2. Average volume measured by Voice Tools also showed a strong correlation with measurements using LingWAVES, this correlation was less strong for highest and especially for lowest volume.

Figure 2 shows the correlation between average voice frequency measured using LingWAVES and using Voice Tools or Voice Pitch Analyzer. This demonstrates that Voice Tools slightly overestimates the average voice frequency in the higher voice frequencies (>150 Hz). Median deviation was 2 Hz (range -4 to 20). On the other hand, Voice Pitch Analyzer slightly underestimates voice frequency in the higher frequencies (>210 Hz). Median deviation was -2 Hz (range -16 to 14).

Group differences

Figure 3 shows the lowest, average and highest voice frequency in the different groups during the first measurement with LingWAVES. Average voice frequency in transmen who received testosterone was comparable to that in control men (difference 1 Hz, 95%CI,-8 to 5). However, untreated trans men had a higher average voice frequency (difference 69 Hz, 95%CI,35 to 103) compared to individuals treated with testosterone. Although transwomen had an average voice frequency of 194 Hz (IQR 136 to 217) which was 26 Hz higher than control men (95%CI, 9 to 44), it was still 22 Hz lower than control women (95%CI, -47 to 2). Besides, a wide range in voice frequency was seen in transwomen.

DISCUSSION

The current study aimed to validate the mobile phone applications Voice Tools and Voice Pitch Analyzer with the registration program LingWAVES as the gold standard to measure voice frequency. It showed that both Voice Tools and Voice Pitch Analyzer give an accurate measurement of the voice frequency and therefore can be used for research and clinical purposes.

The two voice measurements performed with LingWAVES were very closely correlated which shows that there is little intrapersonal variability when measuring voice frequency. This is in line with findings by Printz et al who showed no significant differences in frequency between two voice measurements with three to seven days in between.[18] A slightly weaker correlation regarding voice volume, especially lowest volume, was found in the current study which was also described by Printz et al This could be due to a high microphone sensitivity or a variation in voice volume between two consecutive measurements..

TABLE 2.
Pearson's Correlation Coefficients for Correlations Between the Two LingWAVES Measurements as Well as Between LingWAVES and the Voice Applications VoiceTools and Voice Pitch Analyzer

Voice Characteristic	LingWAVES 1 vs. LingWAVES 2	LingWAVES 1 vs. VoiceTools	LingWAVES 2 vs. Voice Pitch Analyzer
Lowest frequency (Hz)	**0.9753** ($P < 0.001$)	**0.9598** ($P < 0.001$)	**0.9636** ($P < 0.001$)
Highest frequency (Hz)	**0.9859** ($P < 0.001$)	**0.9646** ($P < 0.001$)	**0.9653** ($P < 0.001$)
Average frequency (Hz)	**0.9912** ($P < 0.001$)	**0.9940** ($P < 0.001$)	**0.9909** ($P < 0.001$)
Lowest volume (dB)	**0.4694** ($P = 0.001$)	**0.5990** ($P < 0.001$)	-
Highest volume (dB)	**0.8979** ($P < 0.001$)	**0.8691** ($P < 0.001$)	-
Average volume (dB)	**0.9506** ($P < 0.001$)	**0.9467** ($P < 0.001$)	-

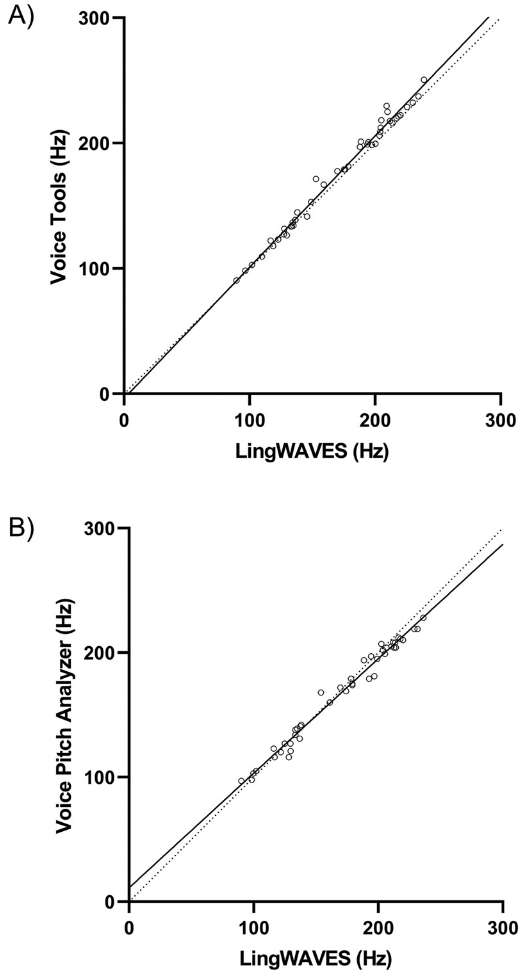

FIGURE 2. Correlation between LingWAVES and (**A**) Voice Tools and (**B**) Voice Pitch analyzer for the average voice frequency (Hz). Line of best fit (solid) and line of perfect correlation (dashed) are drawn. Each measurement is represented as a dot.

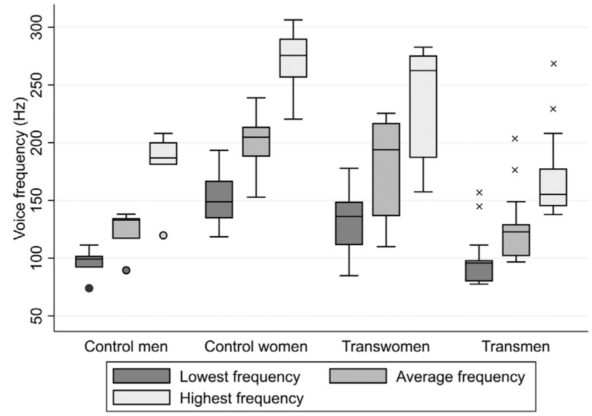

FIGURE 3. Boxplots of the lowest, highest and average voice frequency (Hz) in control men, control women, trans women and trans men in the first measurement with LingWAVES. Boxes represent interquartile range(IQR), whiskers represent 5th to 95th percentile, circles represent outliers and X's represent transmen who have not (yet) started gender affirming hormone treatment.

Average voice frequency, but also lowest and highest pitch, measured using Voice Tools as Voice Pitch Analyzer were highly correlated with measurements using LingWAVES. This is in line with previous studies which showed that other mobile phone applications accurately measured voice frequency when compared to LingWAVES.[12-15]

Voice Tools was shown to overestimate the frequency in higher ranges compared to LingWAVES. Female voices are characterized by higher frequencies.[1,2] Thus, when interpreting female voice frequencies measured with Voice Tools one should be aware of potential overestimation although the deviation did not exceed 20Hz. In contrast, Voice Pitch Analyzer underestimates the voice frequency in higher ranges by a maximum of 16Hz.

The voice frequencies in the different groups of participants showed that trans men who are treated with GAHT and control men have comparable voice frequencies. This is in line with findings from Deuster et al who described that after 52 weeks of treatment no significant difference in speaking fundamental frequency was found between trans men and control men.[19] Although voice frequency in trans women was on average only 22 Hz lower than in control women, a large range in voice frequency was observed in trans women. Their voice frequency of 194 Hz (IQR 136 to 217) was higher than described in a study by Villas-Bôas et al who found a mean voice frequency of 159 Hz in 30 transgender women.[20] This could be explained by the fact that trans women who had had speech therapy or surgery were excluded from the latter study while these therapies are known to significantly increase voice frequency.[9] Since data on voice therapy or surgery were not available in our cohort, the wide range might be explained by the heterogeneity of vocal treatment. Alternatively, it may be due to the fact that some participants may have started treatment in adolescence before completing endogenous puberty and therefore have a higher voice frequency.

This study has strength and limitations. A strength is the fact that voice analysis was performed using both the applications to be validated and LingWAVES simultaneously, in order to compare them using the exact same recording. A limitation is that Voice Pitch Analyzer only reported median voice frequency whereas LingWAVES reported average voice frequency. Another limitation of this study might be the generalization. This study included healthy transgender people and their family, both without other voice problems. However, usability of the apps for other groups such as patients with voice disorders, cannot be derived from this study. Therefore, future research can be performed in other target groups to establish evidence of the broader usability of mobile phone applications to measure voice frequency. Besides, it is uncertain if the results apply to other mobile phones in which the microphone may differ.

The results of this study demonstrate that it is possible to take reliable voice measurements with mobile phone applications. This implicates that voice samples for clinical or research purposes can be made without high costs or the need for specialized equipment. This is ideal for use in low-resource settings and also allows measurements at home, for example enabling people to monitor their progress when doing speech therapy exercises at home. Both Voice Tools and Voice Pitch Analyzer accurately measure voice frequency, but only Voice Tools measures voice volume. Future research is needed to assess the validity of these applications with other mobile phones and in people with voice disorders.

DISCLOSURES

No disclosures.

CONFLICT OF INTEREST

The authors have no conflict of interest.

REFERENCES

1. Latinus M, Taylor MJ. Discriminating male and female voices: differentiating pitch and gender. *Brain Topogr*. 2012;25:194–204.
2. Simpson AP. Phonetic differences between male and female speech. *Lang Linguist Compass*. 2009;3:621–640.
3. Kennedy E, Thibeault SL. Voice-gender incongruence and voice health information-seeking behaviors in the transgender community. *Am J Speech Lang Pathol*. 2020;29:1563–1573.
4. Hembree WC, Cohen-Kettenis PT, Gooren L, et al. Endocrine treatment of gender-dysphoric/gender-incongruent persons: an endocrine society clinical practice guideline. *J Clin Endocrinol Metab*. 2017;102:3869–3903.
5. Deuster D, Di Vincenzo K, Szukaj M, et al. Change of speech fundamental frequency explains the satisfaction with voice in response to testosterone therapy in female-to-male gender dysphoric individuals. *Eur Arch Otorhinolaryngol*. 2016;273:2127–2131.
6. Gray ML, Courey MS. Transgender voice and communication. *Otolaryngol Clin North Am*. 2019;52:713–722.
7. Cosyns M, Van Borsel J, Wierckx K, et al. Voice in female-to-male transsexual persons after long-term androgen therapy. *Laryngoscope*. 2014;124:1409–1414.
8. McNeill EJ. Management of the transgender voice. *J Laryngol Otol*. 2006;120:521–523.
9. Nolan IT, Morrison SD, Arowojolu O, et al. The role of voice therapy and phonosurgery in transgender vocal feminization. *J Craniofac Surg*. 2019;30:1368–1375.
10. Behrman A. Common practices of voice therapists in the evaluation of patients. *J Voice*. 2005;19:454–469.
11. Lee SH, Yeo JO, Choi JI, et al. Local steroid injection via the cricothyroid membrane in patients with a vocal nodule. *Arch Otolaryngol Head Neck Surg*. 2011;137:1011–1016.
12. Maryn Y, Ysenbaert F, Zarowski A, et al. Mobile communication devices, ambient noise, and acoustic voice measures. *J Voice*. 2017;31:248.
13. Munnings AJ. The current state and future possibilities of mobile phone "Voice Analyser" applications, in relation to otorhinolaryngology. *J Voice*. 2020;34:527–532.
14. Grillo EU, Brosious JN, Sorrell SL, et al. Influence of smartphones and software on acoustic voice measures. *Int J Telerehabilitation*. 2016;8:9.
15. Mat Baki M, Wood G, Alston M, et al. Reliability of opera VOX against multidimensional voice program (MDVP). *Clin Otolaryngol*. 2015;40:22–28.
16. Caffier PP, Möller A, Forbes E, et al. The vocal extent measure: development of a novel parameter in voice diagnostics and initial clinical experience. *Biomed Res Int*. 2018;2018: 3836714.
17. Kyriakou K, Theodorou E, Petinou K, et al. Correlation between a self-voice assessment and objective-voice evaluation outcomes in speech language pathology students. *J Voice*. 2022.
18. Printz T, Godballe C, Grøntved Å M. The dual-microphone voice range profile assessment-interrater reliability. *J Voice*. 2021;35:521–529.
19. Deuster D, Matulat P, Knief A, et al. Voice deepening under testosterone treatment in female-to-male gender dysphoric individuals. *Eur Arch Otorhinolaryngol*. 2016;273:959–965.
20. Villas-Bôas AP, Schwarz K, Fontanari AMV, et al. Acoustic measures of Brazilian transgender women's voices: a case-control study. *Front Psychol*. 2021;12: 622526.