Comparing the accuracy of two phonetography systems

Nijkamp, M. [a] BsC, External promoter: Maryn, Y. [b] PhD

[a]Zuyd University of Applied Sciences, Department of Speech and Language Pathology, Heerlen, Netherlands. [b]Department of Otorhinolaryngology and Head & Neck Surgery. GZA Sint-Augustinus, Wilrijk, Belgium. Department of Speech, Languages and Hearing Sciences, Ghent University, Ghent, Belgium. School of Logopedics, Faculty of Psychology and Educational Sciences, Université Catholique de Louvain, Louvain-La-Neuve, Belgium. Department of Speech-Language Therapy, University College Ghent, Ghent, Belgium.

Introduction

The phonetogram shows the estimation of the vocal intensity levels (IL) over the total fundamental frequency (f0) range. A study to compare the accuracy of phonetography systems has never been conducted. In the current study, the accuracy of the lingWAVES phonetography system (WEVOSYS) and the Praat-based (Boersma, 2002) Vox Phonetography (Nijkamp, 2018) are determined and compared for their analysis of synthesized vowels.

Methods

In total, 17 semitones ranging from E2 to E6 were chosen. Vowels were synthesized on these fundamental frequencies with formant frequencies of F1 800 and F2 1300 Hz and a duration of 2 seconds. The sounds were emitted on 40, 50 and 85 dB (at 30 cm from the loudspeaker) to both systems in an anechoic room. Both systems were used according to their protocols. Their output was then compared to the true values.

Results

Results show that the Praat based Vox system makes less measurement error than the lingWAVES system. Raw measurement data is shown in figures 1 and 2, error statistics are shown in table 1.

Measurement error	Vox error (± SD)	lingWAVES error (± SD)	Sig. (t-test)
RMSE f0 (Hz)	**0.098 (± 0.361)**	15.516 (± 37.646)	P = 0.0043**
RMSE f0 (ST)	**0.002 (± 0.006)**	0.469 (± 0.936)	P = 0.0006***
RMSE I (dB)	**1.471 (± 1.433)**	2.258 (± 1.527)	P = 0.0211*
RMSE /C	**1.472 (± 1.434)**	2.727 (± 1.98)	P = 0.0014**

Table 1:RMSE = root mean square error, f0 (Hz) = error in f0 determination expressed in Hertz, f0 (ST) = error in f0 determination expressed in semitones, I (dB) = error in IL determination expressed in decibels, RMSE/C = root mean square error per coordinate, average of the sum of the mean error in semitones and the mean error in decibels. For every row, the smaller systems error is made bold. T-test results show that the differences between the systems are all statistically significant. * = $P < 0.05$, ** = $P < 0.01$, *** = $P < 0.001$.

Conclusion

The Praat based Vox system, given the condition of using an 8 cm microphone to mouth distance and correct intensity level calibration, is more accurate in estimating f0 and IL variables from synthesized vowels than the Lingwaves phonetography system.

Contact Information

Marthè Nijkamp, speech and language therapist at Logopedie Maastricht Oost, Maastricht, Netherlands
Email: nijkamp.marthe@gmail.com

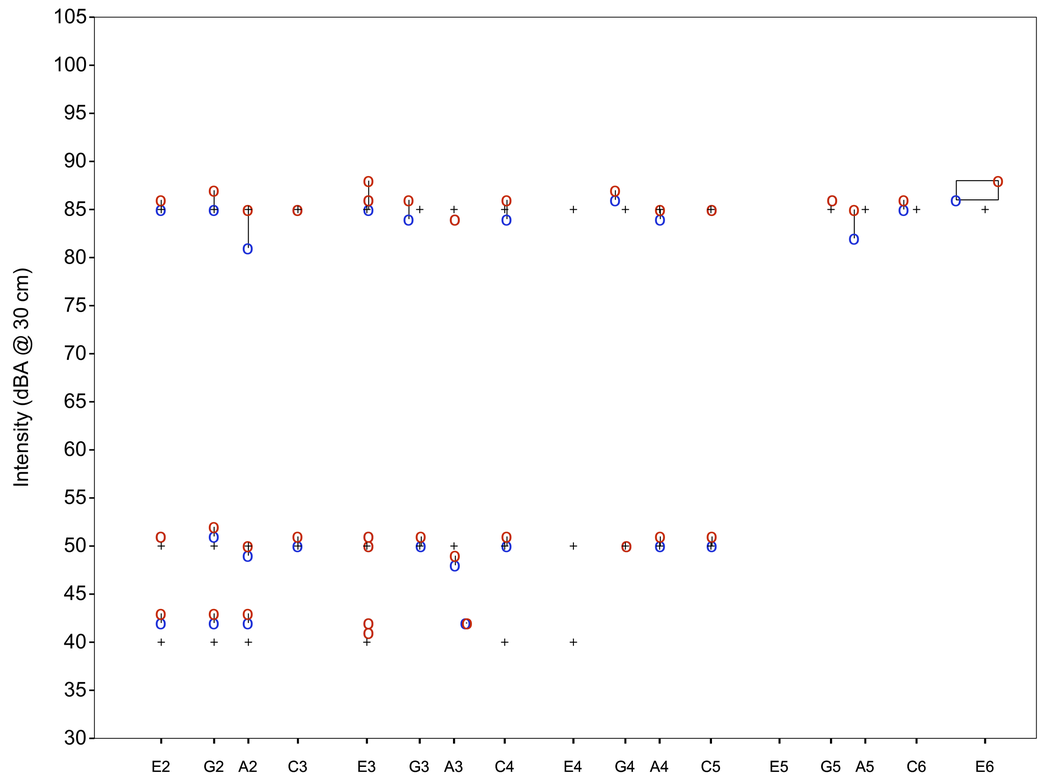

Figure 1:The phonetogram forms visualize the measurements made by Lingwaves. The blue and red dots represents the minimum and maximum IL and f0 measurement respectively. When the system did not respond to the stimulus, no data was drawn for that measurement.

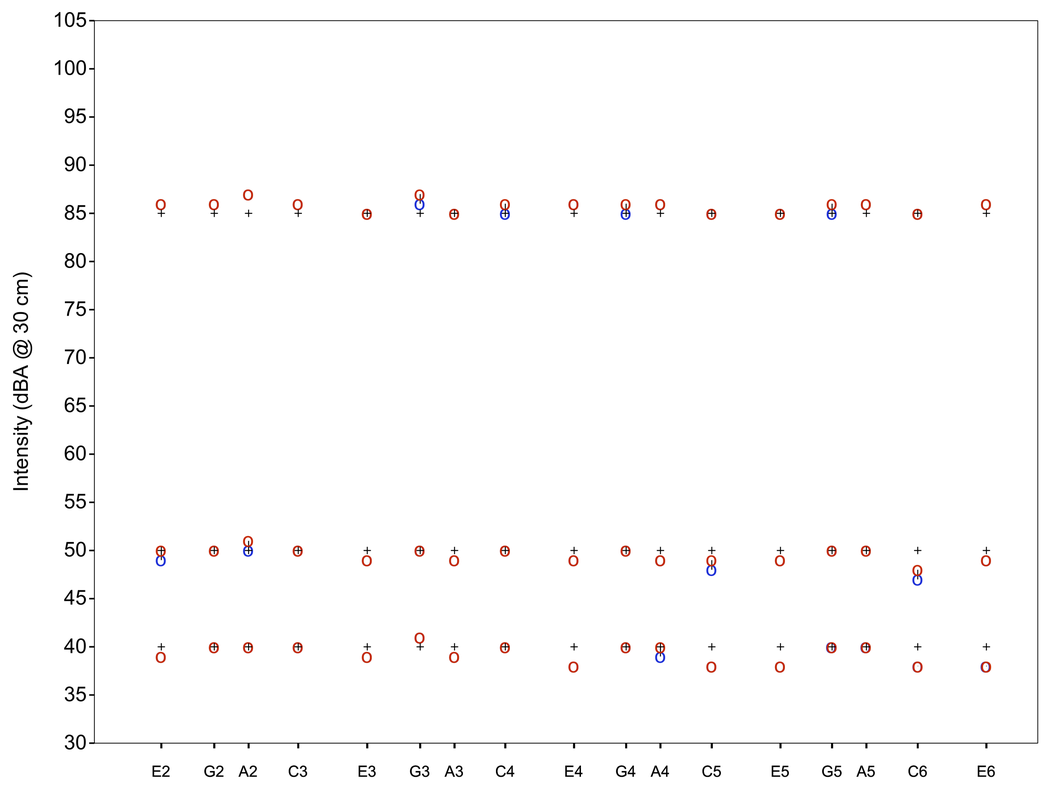

Figure 2:The phonetogram forms visualize the measurements made by Vox. The blue and red dots represents the minimum and maximum IL and f0 measurement respectively. When the system did not respond to the stimulus, no data was drawn for that measurement.

References

Boersma, P. (2002). Praat, a system for doing phonetics by computer.
Nijkamp, M. (2018). Vox Phonetography. Heerlen.
WEVOSYS. (2015). lingWAVES 3: Computer based standardized measurement system for voice and speech assessment, biofeedback and documentation. Global handbook for SLP and Voice Clinic Suites.