

The Intersection of Acoustics and Audience: What Measurement Says About Perceived Voice Acting Quality.

By: TJ Hatch

CSCI 4960: Intro to Research

Professor: Neha Keshan



Motivation



'Objective' quality analysis of a line read has long been thought to be intractable on the basis that perceived quality of a line read is purely subjective. Where drawings can be measured based on how well the underlying forms match 3D space and works of fiction can be judged on grammar and internal consistency, it is hard to label a line read "correct" on any basis other than the presence or absence of microphone peaking, plosives, and audible breaths.

This study is a proof-of-concept of a larger study, designed with the intent to show how we might study voice acting quality and pursue it on more rigorous terms using acoustic analysis. The study aims to answer the question: "how should we approach the study of vocal quality with respect to acoustic analysis?"



Procedure



(Wheatley, Portal 2)

In order to approach this question, we hypothesized a potential method of study and carried it out. To this end, we downloaded five voice lines spoken by Stephen Merchant's Wheatley in Portal 2. Using qualitative data from a survey we would conduct later, we directed a pedestrian to read the same lines, bringing our total number of analyzable lines to 10. We identified four acoustic properties (cadence, pitch, pitch variance, amplitude variance) and measured all voice lines on those fronts. Then we sent out a four-question qualitative survey to ~15 participants, with two qualitative questions and two quantitative questions. The quantitative questions measured ease of identifying the emotion in the speaker's voice and how genuine the listener believed the speaker to sound.

- During the study, we were vigilant for areas of improvement and new questions that were raised while going through the motions. While the findings of the study are interesting themselves, those areas of improvement are the true fruit of this study; they show how the concept of the study can grow and evolve. Areas of expansion include:
- More sophisticated surveys (quality measurements)
 - More sophisticated acoustic measurements (pitch profiles, fundamental frequency normalization, etc)
 - Larger body of voice clips from more diverse actors
 - Acoustic properties of emphasis, etc

Measurement and Survey Results

The professional voice actor overall displayed a higher cadence and pitch variance, but a lower amplitude variance. Differences in base pitch between the two actors is likely due to natural vocal differences. One interesting thing to note is that the range of the professional's measurements was overall higher for each measurement except pitch variance.

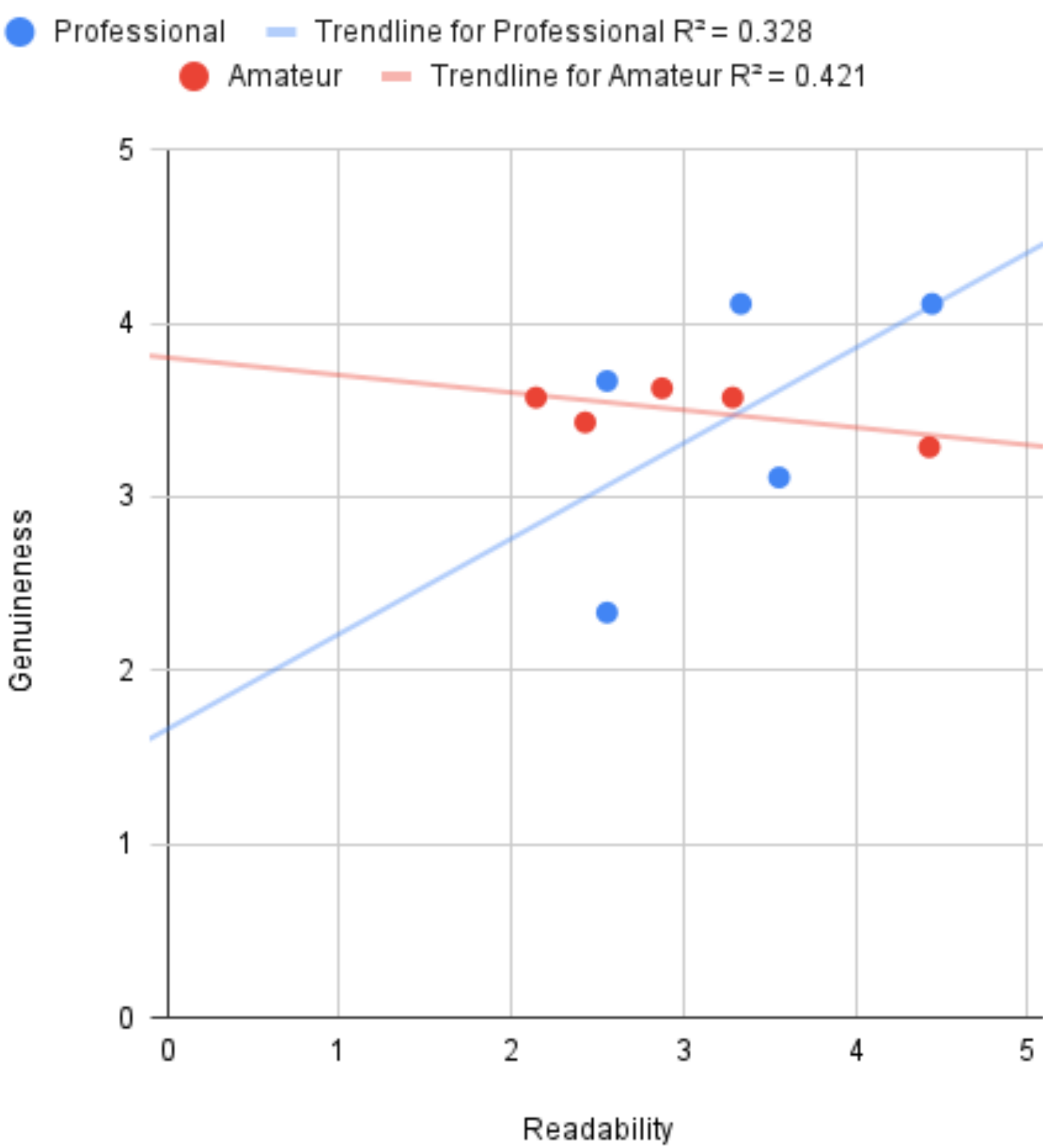
Professional	1	2	3	4	5
Cadence (words per second)	4.054	3.77	3.339	4.319	3.344
Average Pitch (Hz)	279.25	251.3333333	152.4117647	161.0769231	232.2
Pitch Variance (Hz)	85	150	346	176	170
Amplitude Variance (dB)	-4.8259	-8.706	-4.17	-6.215	-0.296
Amateur	6	7	8	9	10
Cadence (words per second)	3.581	2.683	3	2.764	2.668
Average Pitch (Hz)	216.6	189.625	169.4545455	135.1428571	157.4
Pitch Variance (Hz)	385	135	119	63	78
Amplitude Variance (dB)	-10.16	-11.098	-7.804	-13.739	-3.255

These are the average scores for each voice line on both quantitative questions. Each dot is aggregated from ~8 data points. The amateur voice actor is rated much more consistently on readability, hovering around 3.5/5 for all lines. Every genuineness score for the amateur fell between the genuineness scores for the professional's 2nd- and 3rd-most genuine lines. Even with only 5 data points, an F-test on this data yields a P-value well under .05. This is likely the most statistically significant finding in the study.

It is notable that no line averaged below 2 on any question. For both the professional and the amateur, correlations between genuineness and readability are weak, and require more data points to understand. The professional's average scores across all lines are slightly higher for both readability and genuineness, but these differences are slight and not statistically significant.

Neither of these measures are a direct indicator of acting quality. The professional's high variance on genuineness scores (and readability scores, for that matter) raise the possibility that other methods may be more useful for measuring acting quality directly.

Line-By-Line Readability and Genuineness



Thank you to Tim for recording your voice, to my survey participants for your help, and to my classmates and TAs for your feedback.

