Research Paper: Duration of Stuttered Syllables Measured by “Computerized Scoring of the Stuttering Severity (CSSS)” and “Pratt”

Hossein Rezai1, Neda Tahmasebi1, Peyman Zamani1, Mohammad Hossein Haghhighzadeh1, Masume Afshani3*, Ferdous Tahzibi3, Ali Heydari3

1. Musculoskeletal Rehabilitation Research Center, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.
2. Department of Biostatistics, School of Health, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.
3. Student Research Committee, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.

* Corresponding Author:
Masume Afshani, MSc.
Address: Student Research Committee, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran.
Tel: +98 (61) 33743101
E-mail: ma_afshani@yahoo.com

Objectives: This investigation aimed to compare the performance of two software packages, namely “Computerized Scoring of the Stuttering Severity (CSSS)” and “Pratt,” in calculating the mean duration of stuttered syllables.

Methods: In this descriptive-analytical study, 35 eligible stuttering subjects (26 male and 9 female), aged between 18 and 42 (m=26.23±6.02) were selected via the non-probability sampling method to enter the study. Spontaneous speech samples of subjects with stuttering were recorded with the help of a video camera. Two separate tasks were used to calculate the Mean Duration of the Three Longest Stuttering Events (MDTLSE) by applying the second version of the CSSS software and the 5.3.78 version of the Praat software. In the first task, MDTLSE was measured 10 times for a subject with mild stuttering and a subject with severe stuttering. In the second task, MDTLSE was measured just one time for each stuttering participant, and comparison was performed by paired t-test using the SPSS Version 22.0 computer software.

Results: In the first task, in which Praat and CSSS-2 were used for the calculation of MDTLSE (10 times), the minimum and maximum obtained values were found to differ by 0.007 seconds (7 milliseconds) and 0.2 seconds (200 milliseconds), respectively. In the second task, in which MDTLSE was calculated with the software CSSS-2 and Praat, the differences were 2.34±2.17 and 3.02±2.98, respectively, that were found to be statistically significant (P=0.025).

Discussion: Higher reliability and replicability of duration values calculated by Praat software indicates that this software can be applied for more precise determination of duration of stuttered syllables.

ABSTRACT

Objectives: This investigation aimed to compare the performance of two software packages, namely “Computerized Scoring of the Stuttering Severity (CSSS)” and “Pratt,” in calculating the mean duration of stuttered syllables.

Methods: In this descriptive-analytical study, 35 eligible stuttering subjects (26 male and 9 female), aged between 18 and 42 (m=26.23±6.02) were selected via the non-probability sampling method to enter the study. Spontaneous speech samples of subjects with stuttering were recorded with the help of a video camera. Two separate tasks were used to calculate the Mean Duration of the Three Longest Stuttering Events (MDTLSE) by applying the second version of the CSSS software and the 5.3.78 version of the Praat software. In the first task, MDTLSE was measured 10 times for a subject with mild stuttering and a subject with severe stuttering. In the second task, MDTLSE was measured just one time for each stuttering participant, and comparison was performed by paired t-test using the SPSS Version 22.0 computer software.

Results: In the first task, in which Praat and CSSS-2 were used for the calculation of MDTLSE (10 times), the minimum and maximum obtained values were found to differ by 0.007 seconds (7 milliseconds) and 0.2 seconds (200 milliseconds), respectively. In the second task, in which MDTLSE was calculated with the software CSSS-2 and Praat, the differences were 2.34±2.17 and 3.02±2.98, respectively, that were found to be statistically significant (P=0.025).

Discussion: Higher reliability and replicability of duration values calculated by Praat software indicates that this software can be applied for more precise determination of duration of stuttered syllables.
1. Introduction

Stuttering usually begins in childhood and if left untreated exerts a negative influence on the social, emotional, and mental health of the subject [1]. The necessity for making diagnostic evaluation [2], documentation and comparative analysis of the therapeutic outcomes during and after treatment [3], and developing evidence-based treatment approaches [4] are the most important sources of motivation for many investigators to develop reliable scales for measuring stuttering severity. Some of these scales deal with emotional and psychological impacts of stuttering while others deal with audible and visible characteristics of stuttering (e.g. stuttering like disfluencies and physical behaviors accompanying stuttering) [5].

Given the availability of a vast variety of procedures and tools for measuring stuttering behaviors, numerous studies on estimating the reliability of stuttering measurement methods were conducted in the past two decades [6-8]. The findings obtained from these studies indicated that application of different assessment procedures for the assessment of stuttering in a single person may cause variations in the stuttering frequency score thus calculated. The same was shown in two studies published by Rousseau et al. (2008) and O’Brien et al. (2015). These studies evaluated the differences between audio-only and audio-visual procedures of stuttering frequency measurement in preschool children and adults, respectively [9, 10].

In the audio-only procedure, the examiner has access to only the audio format of the subject’s speech sample. On the other hand, in the audio-visual procedure the subject’s speech samples are videotaped, due to which both audio and visual information are available to the examiner. Rousseau et al. (2008) found that the recorded stuttering frequency was significantly lower (by around 20%) when it is measured on the basis of audio-only recordings [9] Conversely, in the study conducted by O’Brien et al. (2015), it was found that when samples were analyzed through the audio-visual mode, a mean 18% increase in the percentage of Stuttered Syllables (%SS) as compared with that of the audio-only procedure was observed [10].

Versions 3 and 4 of the Stuttering Severity Instrument (SSI-3 and SSI-4) are the most common scales for assessing overt stuttering symptoms [5, 11]. In both versions of the SSI, the overall severity score of stuttering is measured by combining the scores of percentages of Stuttered Syllables (%SS), Mean Duration of the Three Longest Stuttering Events (MDTLSE), and Physical Concomitants (PC), which are the face and body movements made during speech sampling of the subjects.

Furthermore, counting the number of fluent and stuttered syllables as well as measuring the MDTLSE are performed by using different instruments, including electronic devices and/or software programs. Stopwatch and Praat [12] are most often used to measure the duration of stuttered syllables. True-Talk [13], Disfluency Frequency Counter [14], and Stuttering Measurement System (SMS) [15] are also used to calculate the number of stuttered and fluent syllables.

Stopwatch, preferably digital, is the most commonly used tool for measuring the duration of stuttered syllables. While watching the videotaped speech sample of the subject, the clinician can start the stopwatch as the stutter begins and stops it when it ends so as to measure the duration of stutter [2]. However, the stopwatch cannot be used for measuring the duration of stuttering events that last less than 1 second [16].

Computerized Scoring of Stuttering Severity Version 2 (CSSS-2) is a commercial software developed by Bakker and Riley (2009) and is under the copyright protection by PRO-ED Inc. [17]. Some studies used this software to determine the participant’s stuttering severity [18, 19]. As shown in Figure 1, CSSS-2 can be applied for counting the syllables (fluent and stuttered), calculating the percentage of stuttered syllables, determining MDTLSE, and measuring speech rate on the basis of the number of syllables per minute (SPM). However, since, CSSS-2...
cannot play the speech sample itself, the examiner needs to use a media player to play the sample. After playing the audio or video file of the speech sample, the experimenter will have to return to the CSSS-2 window and press the keyboard’s space button while listening to the first and last word of the speech sample.

CSSS-2 automatically calculates the duration of the played speech sample by measuring the interval between the pressing of the space button. On the other hand, the left and right buttons of the mouse are used to count the fluent and stuttered syllables, respectively. The experimenter can click the left button of the mouse once for each fluent syllable uttered by the stuttering subject [11]. This way, the total number of fluent syllables would equal the total number of times the left button on the mouse has been clicked. Similarly, the total number of stuttered syllables is calculated by clicking the right button of the mouse, once for each stuttered syllable.

Furthermore, the software automatically reports the percent syllables stuttered (%SS) by dividing the number of stuttered syllables by the total number of syllables in speech sample, which was 250 in this study.

To measure the duration of a stuttered syllable, the right button of the mouse is held down over the length of the stutter. Stuttering events with durations of 0.1 second or more can be measured by CSSS-2. After doing so for all the stuttered syllables, the software automatically detects the 3 longest stuttering events. CSSS-2, calculates MDTLSE by dividing the sum of duration values for three longest stuttering events by 3. Finally, the total number of uttered syllables (fluent/stuttered) is divided by the total duration of the played speech sample and multiplied by 60 to calculate the speech rate of the subject as words per minute (SPM).

Praat is a software program developed by Paul Boersma and David Weenink at the University of Amsterdam. It analyzes the speech sample and expresses them as waveforms and spectrograms that are displayed on a screen. Additionally, Praat allows measuring the duration of the speech segments in any length (syllable, word, phrase, sentence, and paragraph). This can be done by the following steps: A) Placing the cursor at the beginning of a given speech segment; B) Pressing and holding the left button of the mouse; and C) Moving horizontally to the right until the end of the speech segment; and D) Clicking the “sel” icon at the bottom left of the “view and edit” window. Following these steps (as shown in Figure 2), Praat can automatically calculate the duration of the selected speech segment within a second or sometimes within a few milliseconds [20].

Search and review of related electronic literature revealed that in almost all the previous stuttering measurement studies, out of the three components required for obtaining an SSI score, %SS was most often chosen as the main variable [9, 10, 21-25]. This meant that MDTLSE or PC was excluded in most stuttering measurement studies. It was also noted that concerns regarding objective scoring [26], ambiguity of assessment procedures [27], and finally having the lowest reliability compared with %SS or MDTLSE [28], have made...
the PC as the least studied component of SSI. Evidently, there are a variety of applications and tools available for measuring the duration of stuttered syllables. The present study was aimed to examine the similarities and differences in the mean duration of stuttered syllables as obtained by CSSS-2 and Praat so as to determine whether these two different tools provide the clinician with the same values or not. Accordingly, these two research questions were put forth in the present study: Which software provides more constant values in multiple measurements of MDTLSE?; and Are there any significant differences between the mean MDTLSE as measured by Praat and CSSS-2?

2. Methods

Participants

The study presented in here is an analytical-descriptive cross-sectional study. The non-probability sampling method was used for the selection of participants for the study from among the speech therapy clients visiting the clinics in Isfahan city. Following the specifications of a pilot study and using the below formula, the study sample size was determined to be 29 subjects with stuttering.

\[
\begin{align*}
\text{n} & = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \times (Sd)^2}{(d/2)^2} \\
& = \frac{(2.57 + 1.65)^2 \times (1.405)^2}{(1.1)^2} = 35.16 \times 29
\end{align*}
\]

To be more precise, we included 35 subjects with stuttering who met the following inclusion criteria: scored 18 or higher on the Stuttering Severity Instrument – 4 (SSI4), not using speech therapy interventions up to 3 months before entering the study, having no structural abnormalities of the articulators or any neurological or psychiatric disorders. Also, 35 age- and sex-matched subjects with fluent speech who met the inclusion criteria (having fluent and clear speech, having no structural or functional abnormalities affecting the orofacial structures, and having no type of neurological or psychiatric disorders) also participated in the study as controls.

Recording and measurement procedures

For speech sampling, all the participants were interviewed by the second author in a quiet 3 by 4 room with sufficient lighting. Sitting in a stacking chair, participants were requested to talk for 5 minutes about any topic of their interest. A Sony HDR PJ760v Camcorder attached to a Manfrotto MT055CXPRO3 tripod was used to record the sampling sessions. Since Praat software can only analyze audio files, AVS audio converter software [29] was used to convert the video files to audio files. By focusing on the middle part of each participant’s speech sample, segments including 250 speech syllables were selected and extracted to a WAV audio file with the help of Praat software.

The extracted audio files were used for determining the MDTLSE by using either Praat or CSSS-2. Following the procedures explained in the introduction section, MDTLSE was measured with Praat and CSSS-2. In this study, two tasks were performed for determining the MDTLSE. In the first task, one participant with mild stuttering and one with severe stuttering problem (participants 17 and 12, respectively) were randomly selected from all the participants with stuttering. Using Praat and CSSS-2, MDTLSE was measured 10 times for each of these two selected subjects. This task was performed with the intention to find out which of the two used software show higher reliability for multiple MDTLSE measurement. It is obvious that the more constant the values obtained in multiple measurements, the more reliable is the software. In the second task, the MDTLSE was measured just one time for each stuttering participant with the help of Praat and CSSS-2.

Data analysis

The SPSS version 18 windows statistical package and paired t test [30] were used to compare and analyze the data obtained from the two software.

3. Results

35 subjects with stuttering (26 males and 9 females) aged between 18 and 42 (26.23±6.02) and 35 age- and sex-matched subjects with fluent speech participated in this descriptive-analytical study. Description of speech fluency characteristics for subjects with and without stuttering is presented in Table 1. Since subjects without stuttering had no stuttering-like disfluencies (SLDs) during conversation, analytic comparison of MDTLSE was only performed for subjects with stuttering. Values obtained by 10 times measurement of MDTLSE with Praat and CSSS-2 are reported in Table 2 for one participant with mild stuttering and one with severe stuttering.

Praat was used to measure the MDTLSE10 times for a subject with mild stuttering. The results obtained yielded values varying from 0.878 to 0.885 seconds. However, performing the same task with CSSS-2 yielded values ranging from 0.7 to 0.9 seconds. Praat and CSSS-2 measurement of MDTLSE (10 times) for a subject with severe stuttering yielded values ranging from 6.275 to...
6.285 seconds and 5.9 to 6.6 seconds, respectively. It was found that MDTLSE values measured by Praat were significantly higher than that of the values measured by CSSS-2 (P=0.025).

4. Discussion

MDTLSE is one of three main factors determining the total severity score in SSI 3 & 4. It is also thought to contribute to stuttering severity estimation in children and adults [11]. This study aimed to compare the performance efficiency of CSSS-2 and Praat in calculating MDTLSE. In the first task, the values obtained by Praat and CSSS-2 measurement revealed that the minimum and maximum values obtained for the subject with mild stuttering differed by 0.007 seconds (7 milliseconds) and 0.2 seconds (200 milliseconds), respectively. For the subject with severe stuttering, these values differed by 0.01 seconds (10 milliseconds) and 0.7 seconds (700 milliseconds), respectively. The findings of the first task showed that regardless of stuttering severity, in multiple measurements of MDTLSE, minimum and maximum values obtained by CSSS-2, ranged higher than those obtained by Praat. In other words, it seems that Praat provides more reliable values in multiple measurements of MDTLSE.

Table 1. Description of stuttered syllables during conversation in subjects with and without stuttering

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Mean (SD)</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without stuttering</td>
<td>0(0.0)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>With stuttering</td>
<td>11.78(7.58)</td>
<td>3.14</td>
<td>37.89</td>
</tr>
</tbody>
</table>

Table 2. Comparison of MDTLSE as measured by Praat and CSSS for two subjects with mild and severe stuttering

<table>
<thead>
<tr>
<th>Times of Measurement</th>
<th>Subject With Mild Stuttering</th>
<th>Subject With Severe Stuttering</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Praat (Seconds)</td>
<td>CSSS (Seconds)</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>0.880</td>
<td>0.9</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td>0.882</td>
<td>0.9</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
<td>0.878</td>
<td>0.7</td>
</tr>
<tr>
<td>4&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.878</td>
<td>0.7</td>
</tr>
<tr>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.882</td>
<td>0.8</td>
</tr>
<tr>
<td>6&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.882</td>
<td>0.8</td>
</tr>
<tr>
<td>7&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.885</td>
<td>0.7</td>
</tr>
<tr>
<td>8&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.885</td>
<td>0.7</td>
</tr>
<tr>
<td>9&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.878</td>
<td>0.8</td>
</tr>
<tr>
<td>10&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0.882</td>
<td>0.8</td>
</tr>
</tbody>
</table>

Table 3. Results of paired t-test for comparison of MDTLSE measured by Praat and CSSS

<table>
<thead>
<tr>
<th>Software</th>
<th>Mean±SD</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Praat</td>
<td>3.02±2.98</td>
<td>0.001</td>
</tr>
<tr>
<td>CSSS</td>
<td>2.36±2.17</td>
<td></td>
</tr>
</tbody>
</table>

In the second task that involved MDTLSE measurement of all the participants with Praat and CSSS-2, the results of paired t-test indicated that MDTLSE measured by Praat was significantly higher than that obtained by CSSS-2 (Table 3). In other words, when the duration of the same stuttered speech segment of all the participants were evaluated with Praat and CSSS-2, it was observed that the differences in the values were significantly different between the two software.

Praat can both play back the speech samples and show the waveform and spectrogram of the speech samples. It was, thus, concluded that while measuring MDTLSE with the help of Praat, the clinician can simultaneously rely on his ears and eyes. So, to some extent, measuring MDTLSE with Praat can be considered as a kind of audiovisual procedure. Owing to the availability of the spectrographic display of the speech sample in Praat, the clinician can measure the duration of the selected speech segments without the interference of reaction time. As found in previous studies, more detailed estimation of stuttering frequency is possible in audiovisual methods as opposed to the audio-only procedures [9, 10].

This can be understood from the processes that use a stopwatch to measure MDTLSE through CSSS-2. In such conditions, the clinician relies just on his ears, due to which the data obtained may partly be influenced by the clinician’s reaction time. Since, the mean auditory reaction time is around 248.61±12.84 milliseconds [31], one cannot expect clinicians to exactly report the MDTLSE values by CSSS-2. Clinician’s delay in holding or realizing the left button of the mouse may also be responsible for higher variability observed in maximum and minimum values obtained by CSSS-2. However, Guitar states that “any delays in starting the stopwatch at the beginning of stutters are compensated for by similar delays when you stop it at the end” [2]. Considering these factors the variability in clinicians’ auditory reaction times may not completely explain the differences observed in the MDTLSE values obtained by CSSS-2.

Using the keywords stuttering, stutterer(s), stutter, disfluency, disfluencies, duration, durational, Praat, and CSSS-2, through an electronic search on PubMed during June-July 2017, we found 7 articles that investigated the duration of SLDs as part of the study [32-36]. Just 1 out of these 7 articles was performed on adults [36] while the rest included children and adolescences as their participants. No exactly similar study to the present investigation was found to have been conducted before. However, in a study by Jani et al. (2013), total spoken syllables, frequency and durations of stuttered syllables were assessed under simultaneous and successive conditions by using Audindex software. In the said study, the researchers simultaneously assessed the total number of spoken syllables, frequency of stuttered syllables, and duration of stuttered syllables while listening to the audio file of each participant’s speech sample in one playback of the speech file. In successive condition, each of the three factors was assessed in three separate sessions [37].

According to Jani et al. [37] audindex is similar to CSSS-2 in applying the procedures for determining stuttering severity available in SSI-4. Thus, the values for duration of stuttered syllables obtained in successive conditions of the said study can be compared with the results obtained in ours. The study conducted by Jani et al. showed that the MDTLSE values measured during simultaneous and successive conditions were 2.37±0.188 and 2.38±0.226, respectively. In our study, the MDTLSE value obtained by CSSS-2 was 2.36±0.17.

While working with CSSS-2, all the clinician needs to do is just pressing the left button of the mouse for each client’s fluent syllable, and holding and releasing the right button of the mouse over the length of the stutter. CSSS-2 automatically performs all the calculations. All the variables measured by CSSS-2 (counting syllables, measuring MDTLSE, and determining speech rate) can be measured by Praat. However, using Praat, the clinician has to do all the measurements manually, which is a more time and energy consuming procedure.

5. Conclusion

Given the larger variability observed in data obtained by CSSS-2 in multiple measurements of MDTLSE, it seems that Praat provides more reliable values. It was also observed that the MDTLSE measured by CSSS-2 was significantly lower than that obtained by Praat. Although CSSS-2 is a software that was specially developed for assessment of stuttering, it seems that Praat can be a more reliable option for determining the duration of stuttered syllables.

Acknowledgments

This study was funded by the Musculoskeletal Rehabilitation Research Center of Ahvaz Jundishapur University of Medical Sciences (grant number= PHT-9512).

Conflict of Interest

All authors certify that this manuscript has neither been published in whole nor in part nor being considered for
publication elsewhere. The authors have no conflicts of interest to declare.

References


