

Taynara Gomes de Araújo<sup>1</sup>  
Thalyta Magalhães Rodrigues<sup>1</sup>  
Renata Maria Moreira Moraes  
Furlan<sup>1</sup>  
Estevam Barbosa de Las Casas<sup>1</sup>  
Andréa Rodrigues Motta<sup>1</sup>

# Reproducibility assessment of an instrument for measuring the axial force of the tongue

## *Avaliação da reprodutibilidade de um instrumento para medição da força axial da língua*

### Keywords

Tongue  
Muscle Strength  
Evaluation  
Equipment and Supplies  
Speech, Language and Hearing  
Sciences

### Descritores

Língua  
Força Muscular  
Avaliação  
Equipamentos e Provisões  
Fonoaudiologia

### ABSTRACT

**Purpose:** Evaluate the reproducibility of Forling, a portable instrument for measuring axial tongue force. **Methods:** Axial force of the tongue was measured in 49 individuals (30 women and 19 men) aged 18-25 years using the Forling portable instrument. Measurements were performed in three days at intervals of 7±2 days. On each day, three 7-second measurements were performed with one-minute intervals between them. The coefficient of variation, Wilcoxon paired test, and intraclass correlation coefficient were used in the statistical analysis of the data. Maximum and mean tongue force values were analyzed, and comparison between them was performed using three approaches: the mean of the three values; the mean of the two highest values; the highest value of each measurement. **Results:** In the analysis of mean tongue force, the coefficient of variation was considered desirable and the intraclass correlation coefficient was acceptable. Significant differences were observed regarding the maximum value between the second and third days, mean of the two highest values and mean of the three values between the first and second days and the second and third days. In the analysis of maximum tongue force, the coefficient of variation and the intraclass correlation coefficient were acceptable. Significant difference was found only in the comparison between the second and third days. **Conclusion:** Good reproducibility of the data obtained with the use of the Forling portable instrument was observed.

### RESUMO

**Objetivo:** Avaliar a reprodutibilidade do Forling, instrumento portátil para medição da força axial da língua. **Método:** Foi realizada a medição da força axial da língua de 49 participantes, 30 mulheres e 19 homens, com idades entre 18 e 25 anos por meio do Forling portátil. As medições foram realizadas em três dias com intervalos de 7±2 dias. Em cada dia, foram realizadas três coletas com duração de 7 segundos e intervalos de 1 minuto. Na análise, empregou-se coeficiente de variação, teste pareado de Wilcoxon e coeficiente de correlação intraclass. Foram analisados os valores de força máxima e média da língua, sendo a comparação dos valores realizada em três abordagens: utilizando-se a média dos três valores; a média dos dois maiores valores; e o maior valor de cada medição. **Resultados:** Na análise da força média, o coeficiente de variação foi considerado desejável e o coeficiente de correlação intraclass aceitável. Houve diferenças significativas entre o máximo valor comparando o segundo e terceiro dias, entre a média dos dois maiores valores e entre a média dos três valores comparando o primeiro e segundo dias, bem como o segundo e terceiro dias. Na análise da força máxima, o coeficiente de variação e o coeficiente de correlação intraclass foram aceitáveis. Somente comparando o segundo e terceiro dias, foi encontrada diferença significativa. **Conclusão:** Observou-se boa reprodutibilidade dos dados obtidos com o uso do Forling portátil.

### Correspondence address:

Andréa Rodrigues Motta  
Faculdade de Medicina, Universidade  
Federal de Minas Gerais – UFMG  
Av. Professor Alfredo Balena, 190,  
sala 251, Santa Efigênia, Belo  
Horizonte (MG), Brasil, CEP: 30130-100.  
E-mail: andreamotta@ufmg.br

Received: September 12, 2017

Accepted: December 20, 2017

Study conducted at Departamento de Fonoaudiologia, Universidade Federal de Minas Gerais – UFMG - Belo Horizonte (MG), Brazil.

<sup>1</sup> Universidade Federal de Minas Gerais – UFMG - Belo Horizonte (MG), Brasil.

**Financial support:** Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – Process nº 538.

**Conflict of interests:** There is a conflict of interest by the authors Renata Maria Moreira Moraes Furlan, Andréa Rodrigues Motta and Estevam Barbosa de Las Casas for being part of the group of inventors who have applied to National Institute of Intellectual Property, Brazil, for the patent of the equipment used in the study.



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## INTRODUCTION

It is essential that force and mobility of the tongue be adequate so that its functions of chewing, swallowing, suction, breathing, and phonoarticulation can be performed harmoniously<sup>(1)</sup>.

Axial force is defined as the force along the axis on which it is exerted. It is thus characterized as a longitudinal force that, in the case of the tongue, refers to protrusion force<sup>(2)</sup>. The protrusion force of the tongue against resistance requires activation of the genioglossus muscle and the intrinsic lingual muscles, and the genioglossus muscle is responsible for providing a stable platform for the intrinsic muscles to develop this force<sup>(3)</sup>. Therefore, assessment of the axial force of the tongue enables verification of a muscle group of extreme relevance in the clinical practice of speech-language pathology - the intrinsic musculature.

In clinical practice, assessment of tongue force is performed perceptually by speech-language therapists, considering that there are few instruments available for this purpose. The Iowa Oral Performance Instrument (IOPI), which provides numerical data on tongue pressure against an air bulb, is the most cited tool in the literature<sup>(4)</sup>. Another instrument, based on the same principle, uses an arrangement of bulbs positioned on the hard palate<sup>(5)</sup>. There are also tools composed of sensors positioned on palatal plates<sup>(6,7)</sup> placed on the palate<sup>(8)</sup> or the teeth<sup>(9)</sup>. The bulbs are disposable and soft, thus hygienic and comfortable, but when they are not fitted to the palate, it is difficult to maintain the reproducibility of their positioning in the oral cavity, because they slip on the surface of the tongue and the tube connecting the bulb to the instrument has no markings indicating its position after the lips are closed<sup>(10)</sup>. Tools that use palatal plates need to be constructed in an individualized way, because each individual has a different size and format of palate<sup>(10)</sup>. In addition, they only enable assessment of force in the cranial direction. However, they are effective in measuring tongue force during its functions, because they enable patients to close their mouths and perform functions almost normally<sup>(10)</sup>. Fitting the sensors directly into the oral cavity is the best way to measure force during functions, but presents difficulties in reproducibility of the sensor fixation point, so that comparisons can be made with the reassessments, and in disinfection of the sensors for use in different patients<sup>(10)</sup>.

Perceptual evaluation is subjective and depends on the examiner's good judgment and experience, whereas instrumental assessment is performed using instruments that provide numerical data on tongue force or pressure, but most tools are not commercially available in Brazil. For this reason, the Biomechanical Engineering Group of the Universidade Federal de Minas Gerais (UFMG) developed an instrument to measure the axial force of the tongue that has already been used in several studies<sup>(1,2,11-17)</sup>. The portable version of the Forling instrument was designed later<sup>(18)</sup>, and is currently undergoing testing.

Analysis of the reproducibility of an instrument should always be considered, and precedes its use in research and clinical practice, because data obtained with methods without assured reproducibility may lead to results that are not representative of the phenomenon and, therefore, not generalizable, threatening the effectiveness and soundness of clinical decision-making<sup>(5)</sup>. Moreover, knowledge about the accuracy or variability of the

data provided by each method enables their consideration during selection of the instrument to be used and interpretation of results<sup>(19)</sup>.

Reproducibility is conceptualized as the degree of agreement between the results of measurements of the same measurand performed under varied conditions. It can be expressed quantitatively depending on the characteristics of the dispersion of results. For an expression of reproducibility to be valid, it is necessary to specify the altered conditions<sup>(20)</sup>, i.e., different locations, days, operators, and measurement systems.

In view of what has been previously exposed, the objective of this study was to assess the reproducibility of the Forling portable instrument for measuring the axial force of the tongue. The hypothesis of this research is that the measurements obtained using the Forling instrument present good reproducibility.

## METHODS

This study was approved by the Research Ethics Committee of the Universidade Federal de Minas Gerais (UFMG) under proc. n° 008/10.

This is a prospective, observational, analytical study conducted with a non-probabilistic sample composed of 49 undergraduate students (30 women and 19 men), aged 18-25 years, enrolled in the Medicine and Speech-Language Pathology courses of the UFMG. All participants signed an Informed Consent Form prior to study commencement.

For recruitment of the sample, an active search was conducted by means of posters and direct invitations at the academic units of the UFMG. Inclusion criteria were as follows: minimum age of 18 years; absence of glossectomy and/or plectomy, tongue paralysis or paresis, and cognitive impairment, identified from an interview conducted prior to data collection. Individuals with clinical classification of tongue force divergent between two examiners and those classified with increased or severely reduced tongue force were excluded from the survey.

Initially, a clinical assessment of the tongue of each participant was performed independently by two previously trained examiners. To evaluate tongue force, participants were requested to protrude the tongue and push it against a wooden spatula positioned vertically a few centimeters away from the lips for seven seconds. Tongue force of the participants was classified as follows: normal, when the musculature was able to protrude the tongue against the firm resistance made by the spatula and maintain the force without trembling and/or deformation; slightly reduced, when the musculature was able to protrude the tongue against the firm resistance made by the spatula and maintain the force, but with slight trembling and upward or downward bending of the apex of the tongue; moderately reduced, when the musculature was able to protrude the tongue against the firm resistance made by the spatula and exert moderate force with trembling and upward or downward bending of the apex of the tongue; severely reduced, when the musculature was weak and only able to withstand slightly against the firm resistance made by the spatula with occurrence of shaking and deformation, being able or not to protrude the tongue out of the oral cavity; increased, when the musculature was able to protrude the tongue

and exert excessive force against the firm resistance made by the spatula<sup>(18)</sup>. Only the participants who had their tongue force classified as normal or slightly reduced by the two examiners were submitted to the instrumental evaluation.

Participants who fit the study profile were assessed using a portable version of the Forling instrument (Figure 1). The instrument is composed of a mouthguard made of thermo-moldable material that adapts to the dental arches of each participant; three pieces made of epoxy, namely, a base part, a fastener, and a drive shaft; and a force-sensing resistor. The base part fits into the center of the mouthguard and provides support for the sensor. The fastener attaches the sensor to the base part. The drive shaft is in contact with the tongue and it is the place where it exerts the force. During measurements, the participant pushes the drive shaft with the tongue with maximum force, pressing the sensor<sup>(18)</sup>. Figure 2 shows the mouthguard with the parts connected and Figure 3 presents the scheme of these fitting parts.

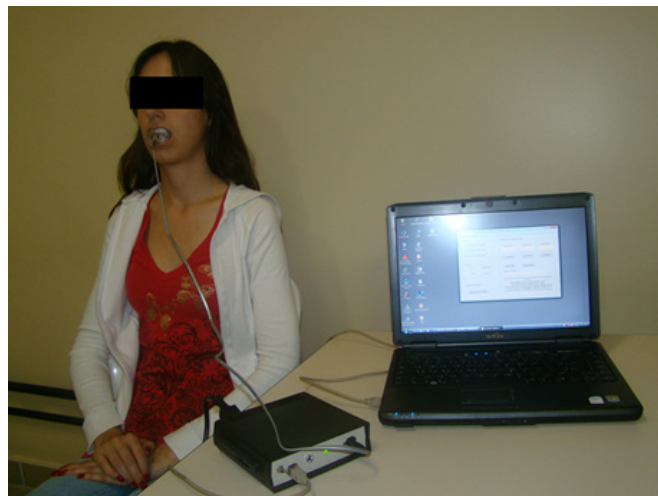
Participants were seated with their backs well supported and their feet flat on the ground. They were instructed about the operation of the instrument and requested to insert the mouthguard into their oral cavity comfortably. The examiner then waited 15 seconds for participants to get used to mouthguard and instructed them to protrude the tongue with the greatest possible force upon request, triggering the device that measures the tongue force for seven seconds<sup>(1,21)</sup>; participants had no access to the test results. All instrumental assessments were conducted by the same previously trained examiner.

Clinical evaluation and the first instrumental assessment were performed on the same day (Day 1), the second instrumental assessment was conducted  $7 \pm 2$  days later (Day 2), and the third instrumental assessment was performed  $7 \pm 2$  days after the second test (Day 3). On each day, three measurements were taken using the Forling instrument with one-minute intervals between them<sup>(1,13,21-24)</sup>. Software developed in the LabView platform showed the force values in newton (N) in real time and recorded the force graphs over time. The same examiner that instructed the participants conducted all the instrumental assessments. Values of maximum and mean force of the tongue were obtained in each measurement. Maximum tongue force refers to the maximum value obtained in the measurement, whereas mean tongue force refers to the arithmetic mean of all the values generated in the measurement.

Random and systematic errors were evaluated by means of differences in the mean value between the days, which may reflect sampling errors and the effect of training. Intra-individual variation was determined by typical error expressed by the coefficient of variation (CV). The CV was assessed for each step, from Day 1 to Day 2 and Day 2 to Day 3, and it was considered acceptable for values  $<10\%$  and desirable for values  $<5\%$ . The CV represents sources of technical and biological errors. Also, the Wilcoxon paired test was applied for each pair. The repeatability of the Forling instrument was investigated by intraclass correlation coefficients (ICC), which was considered acceptable for values  $>0.600$  and desirable for values  $>0.800$ .

Two statistical analyses were performed: the first considered only the values of the mean tongue force of each measurement and the second considered the values of the maximum tongue force

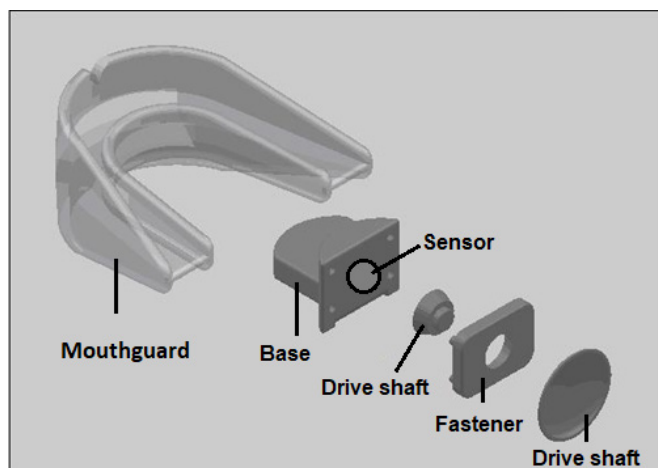
of each measurement. For each of these analyses, comparison between the values (maximum or mean tongue force) was performed using three approaches, as in a previous study that analyzed



**Figure 1.** Individual with a Forling portable instrument fitted in the oral cavity



**Figure 2.** Mouthguard with the parts connected



**Figure 3.** Scheme of the fitting parts of the mouthguard

the Iowa Oral Performance Instrument (IOPI)<sup>(19)</sup>: (1) the mean of the two highest values of instrumental assessment; (2) the highest value of each instrumental assessment; (3) the mean of the three values of each instrumental assessment.

## RESULTS

### Clinical evaluation

Clinical evaluation identified 33 individuals with normal tongue force (67.3%), 21 (63.6%) women and 12 (36.4%) men, and 16 individuals with slightly reduced tongue force (32.7%), seven (43.7%) men and nine (56.3%) women.

### Instrumental assessment: analysis of mean tongue force of participants in each measurement

Initially, the data obtained from the participants in each measurement were analyzed according to their mean value (Table 1). Three measures, one from each measurement, were obtained on each day. These measures were summarized in a single value in three ways: maximum value, mean of the two highest values, and mean of the three values.

In the descriptive analysis, considering the maximum value between the measurements on each day or the mean of the three values, progressive increase in tongue force was observed on the measurement days. When the measurement values were summarized according to the mean of the two highest values found between them, it was verified that the highest value of the measurement occurred on Day 3.

The Wilcoxon test didn't show statistically significant difference between the maximum values measured for Days 1 and 2. (Table 2). The typical error was <5% and the intraclass correlation coefficients (ICC) were acceptable on both days.

When the mean of the two highest values of the measurement on each day was considered, it was verified that the largest mean difference between the measurements occurred from Day 2 to Day 3. The Wilcoxon test showed statistically significant difference between the values obtained. For all pairs of the measurement days, typical error values were <5% - considered desirable. The ICC results for all pairs of days were considered acceptable.

The value for each day was also summarized according to the mean value of the measurements performed. In the reproducibility analysis, statistically significant difference was observed by the Wilcoxon test between the values obtained. The typical error was considered desirable and the ICC results were acceptable on both days.

### Instrumental assessment: analysis of maximum tongue force of participants in each measurement

In a second moment, the data obtained from the participants in each measurement were analyzed according to the maximum value found. Once again, three measures, one from each measurement, were obtained on each day. These measures were summarized in a single value in three ways, as previously exposed, and reproducibility analysis of the test instrument was conducted.

In the descriptive analysis (Table 3), with the data of maximum value of each day, as well as with the data summarized by the mean value of the measurements on each day, an increase in the values was observed throughout the days. When the values were summarized by the mean of the two highest measures obtained, it was verified that the highest mean value occurred on Day 3.

In the analysis of the maximum value on each day, the *p*-value of the Wilcoxon test was low, close to zero, only between Days 2 and 3, indicating difference between the values for these two days. The typical error for Days 1 and 2 was acceptable, and so were the ICC results (Table 4).

**Table 1.** Descriptive analysis of the mean force of the tongue

	Day 1 Mean (Standard deviation)	Day 2 Mean (Standard deviation)	Day 3 Mean (Standard deviation)
Maximum value on each day	13.077 (4.35)	13.753 (4.235)	14.962 (3.976)
Mean of the two highest values on each day	12.443 (4.182)	13.227 (4.019)	14.313 (3.848)
Mean of three values on each day	11.652 (3.893)	12.674 (3.824)	13.843 (3.809)

Tongue force values shown in Newton (N)

**Table 2.** Reproducibility analysis of the mean force of the tongue

	Change in mean	95% CI	<i>p</i> -value (Wilcoxon)	CV Typical error	95% CI	CV Mean	95% CI	ICC	95% CI
Maximum value on each day									
Day 1-2	0.676	(-0.155; 1.507)	0.156	0.030	(0.025; 0.038)	0.023	(0.000; 0.066)	0.766	(0.620; 0.861)
Day 2-3	1.208	(0.389; 2.027)	0.002	0.017	(0.014; 0.021)			0.726	(0.561; 0.836)
Mean of the two highest values on each day									
Day 1-2	0.784	(-0.003; 1.570)	0.049	0.025	(0.021; 0.031)	0.021	(0.000; 0.044)	0.765	(0.618; 0.860)
Day 2-3	1.086	(0.309; 1.863)	0.002	0.018	(0.015; 0.022)			0.735	(0.574; 0.841)
Mean of three values on each day									
Day 1-2	1.022	(0.306; 1.738)	0.006	0.017	(0.014; 0.022)	0.016	(0.010; 0.022)	0.764	(0.617; 0.860)
Day 2-3	1.169	(0.306; 1.738)	0.001	0.015	(0.013; 0.019)			0.744	(0.586; 0.847)

**Caption:** CI = confidence interval; CV = coefficient of variation; ICC = interclass correlation coefficient

**Table 3.** Descriptive analysis of the maximum force of the tongue

	Day 1 Mean (Standard deviation)	Day 2 Mean (Standard deviation)	Day 3 Mean (Standard deviation)
Maximum value on each day	18.939 (6.66)	18.958(5.703)	20.995 (5.684)
Mean of the two highest values on each day	17.775(6.09)	18.217 (5.448)	20.08 (5.413)
Mean of three values on each day	16.818 (5.72)	17.527 (5.215)	19.258 (5.219)

Tongue force values shown in Newton (N)

**Table 4.** Reproducibility analysis of the maximum force of the tongue

	Change in mean	95% CI	p-value (Wilcoxon)	CV Typical error	95% CI	CV Mean	95% CI	ICC	95% CI
Maximum value on each day									
Day 1-2	-0.363	(-1.684; 0.958)	0.542	0.088	(0.073; 0.111)	0.050	(0.000; 0.292)	0.666	(0.476; 0.797)
Day 2-3	-2.037	(-3.021; -1.053)	0.000	0.012	(0.010; 0.015)			0.766	(0.619; 0.861)
Mean of the two highest values on each day									
Day 1-2	0.442	(-0.826; 1.709)	0.636	0.071	(0.059; 0.088)	0.042	(0.000; 0.226)	0.711	(0.539; 0.826)
Day 2-3	1.863	(0.936; 2.790)	0.001	0.012	(0.010; 0.015)			0.774	(0.632; 0.866)
Mean of three values on each day									
Day 1-2	0.708	(-0.457; 1.872)	0.424	0.040	(0.034; 0.051)	0.026	(0.000; 0.116)	0.723	(0.556; 0.833)
Day 2-3	1.732	(0.872; 2.592)	<0.001	0.012	(0.010; 0.015)			0.789	(0.654; 0.875)

**Caption:** CI = confidence interval; CV = coefficient of variation; ICC = interclass correlation coefficient

In the analysis of the mean of the two highest values of the measurements on each day, the largest difference was verified between Day 2 and Day 3. The typical error was considered desirable between Days 2 and 3 and acceptable between Days 1 and 2. ICC results were acceptable for both comparisons.

Regarding analysis of the mean of the three measurements on each day, no statistically significant difference was observed from Day 1 to Day 2 by the Wilcoxon test. It was also verified that the typical error was desirable and both ICC results were acceptable, thus showing good reproducibility.

## DISCUSSION

Reliability of an instrument or test can be determined by the agreement of the results provided by different examiners (interrater reliability) or at different moments (intra-examiner or test-retest)<sup>(25)</sup>. Intra-examiner reproducibility was analyzed in the present study. An option was made for time variation, and three measurements were taken at different times at intervals of 7±2 days, with no changes in the test conditions, considering that this would be the greatest applicability of an instrument of this nature in the area of Orofacial Motricity. A two-day variation was allowed in the instrumental assessments in order to avoid study withdrawal by the participants because of difficulty in schedule. However, assessment conditions were maintained on the three days, avoiding interference with the results.

All descriptive analyses show an increase in tongue force values throughout the measurement days. As a consequence, differences in the mean values observed were large, especially between Day 2 and Day 3, reflecting in the *p*-value (Wilcoxon test), which indicated statistically significant differences between the measures. This might have occurred because of the training

effect, that is, the individuals evaluated were improving their results due to familiarity with the measurement process.

Analysis of the mean tongue force data shows that, regardless of the type of analysis, all typical error values were considered desirable (<5%), which indicates that distribution of the differences in the measurement values of the pairs of days was homogeneous. In addition, the ICC results remained within acceptable range in all analyses, showing good reliability. Nevertheless, the Wilcoxon test indicated statistically significant differences in the maximum value analysis (in the Day 2-3 pair), in the analysis of the mean of the two highest values (in the pairs Day 1-2 and Day 2-3), and in the analysis of the mean of the three values (in the pairs Day 1-2 and Day 2-3). Therefore, the mean tongue force values presented good reproducibility in two of the three assessment methods employed.

In the analysis of maximum tongue force, all typical error values were considered acceptable, and the Day 2-3 pair showed typical error values considered desirable by the three methods employed. With respect to ICC, the results obtained remained within acceptable range in the three types of analysis, also demonstrating good reliability.

Regarding the Wilcoxon test, unlike what was observed in the analysis of the mean tongue force, statistically significant difference was found only in the Day 2-3 pair, and the Day 1-2 pair indicated no significant difference by the three methods used. Thus, just like the mean tongue force values, the maximum tongue force values showed good reproducibility in two of the three assessment methods conducted.

These results indicate that maximum tongue force seems to be a pattern with smaller variation than mean tongue force, as reported in a previous study<sup>(12)</sup>, and the best way to summarize it was using the mean of the three measures.

Results of this study are in agreement with those obtained in other surveys that addressed assessment methods for measuring tongue force and found coefficient of variation values between 0.056 and 0.093<sup>(19)</sup> and between 0.014 and 0.070<sup>(26)</sup>, although they have employed a different instrument - the Iowa Oral Performance Instrument (IOPI), as well as different methodologies for data analysis. Both instruments, Forling and IOPI, propose the assessment of tongue force, but the Forling uses a force sensor and returns data in unit of force (N), and the IOPI uses a pressure sensor and provides data in unit of pressure (kPa); consequently, the measures generated by the instruments cannot be compared. Furthermore, the Forling requires a protrusion movement, whereas the IOPI requires participants to compress the tongue against the hard palate.

All the coefficients of variation found were considered desirable or acceptable, demonstrating homogeneity of the values and confirming the good reproducibility of the data generated by the Forling portable instrument.

However, typical error is the most important analysis for the reliability of the measures, because it indicates the variability of the data from individuals between different assessments<sup>(19)</sup>. Therefore, if this parameter is considered, it can be stated that the instrument is reproducible.

In this research, a tongue force application time of seven seconds was used, considering that individuals with normal tongue force reach maximum force within the first seconds of contraction, but individuals with reduced tongue force need a longer time<sup>(27)</sup>. Some previous surveys have also used the time of 7 seconds<sup>(1,21)</sup>, whereas others have used a shorter time, between two and three seconds<sup>(19,22,26,28)</sup>, but these studies that used shorter contraction times evaluated only maximum tongue force, whereas the present study also assessed mean tongue force. In order not to obtain results with variations caused by muscle fatigue, a rest period was observed between measurements; as in previous studies, this time was one minute<sup>(1,13,21-24)</sup>.

In addition to the good reproducibility of the data, it is worth noting that the software that accompanies the Forling instrument is easy to use, enabling modification of the data recording time and data visualization during measurement, allowing the instrument to be used in different ways and without need of extensive training. The ranges of mean and maximum tongue force values found in this study are consistent with those reported by other authors during assessment of protrusion tasks in healthy adults. A study conducted with adults aged 20-44 years found maximum tongue force of 16.1±4.7 N and mean tongue force of 11.6±3.4 N<sup>(16)</sup>, whereas a survey conducted with men aged 22-39 years obtained maximum tongue force of 24.3±6.7 N<sup>(29)</sup>.

Due to the training factor observed in this study, it is advisable that examiners provide participants with longer familiarization time with the instrument. If participants are assessed in a single session prior to treatment, it should be considered that there will be a gain in tongue force relative to the training, previously or concomitantly with the treatment gain. Adams et al.<sup>(19)</sup> suggested that several measurements be performed in the first session so that patients can become familiarized with the instrument.

Differences in axial tongue force are expected throughout life, but significant variations are not observed in groups of young

adults, as the participants of this study, only in individuals over 60 years of age<sup>(2)</sup>. Some studies have also shown differences in tongue force between men and women, and the values found for males were generally higher than those observed for females<sup>(22,30)</sup>. Despite this difference, reproducibility results are not affected, because comparisons for this type of analysis are conducted with the individuals themselves. For these reasons, analysis stratified by age groups or gender were not performed in this study.

This study was conducted with a small sample and did not include major orofacial myofunctional impairments. Clinical evaluation showed that most participants had normal tongue force, but 32.7% presented slightly reduced tongue force. A previous study<sup>(16)</sup> performed clinical evaluation of the tongue in young adults and found a larger number of individuals (62.5%) with some impairment in tongue force in the task of protrusion against resistance; however, most of them presented weakness in the apex or anterior third of the tongue, and only 4.2% had slightly reduced tongue force. As clinical evaluation is subjective, differences between surveys are expected. The authors justified that minimal changes in tongue force were considered during the qualitative classification, thus a large number of impaired individuals were observed<sup>(16)</sup>. The clinical evaluation of tongue force in the present study was conducted as an inclusion criterion, so that it was possible to initially understand the behavior of the instrument, with respect to reproducibility, in individuals without significant tongue force impairments.

It is believed that better results could be obtained if the study sample were larger and there were more days of measurement in order to mitigate the training effect. It is therefore suggested that further longitudinal studies be conducted with larger samples, longer measurement periods, and with interrater reproducibility assessment.

It was also possible to observe that the type of tongue force chosen (maximum or mean) and the different types of analysis led to different results; therefore, it is important to verify and carefully choose the methodology and the statistical analysis model in order to avoid any bias.

Therefore, it is expected that the Forling portable instrument be used in the research and clinical practice of speech-language pathologists and other professionals who work in the interface of the Orofacial Motricity area, considering that quantitative assessment enables observation of patient evolution throughout therapy in a simple and reliable way.

## CONCLUSION

In the present study, very good reproducibility of the data obtained using the Forling portable instrument was observed, with most parameters assessed - typical error, mean coefficient of variation, confidence interval, and intraclass correlation coefficient - presenting acceptable and desirable results, enabling the use of this instrument both in clinical practice and in other studies.

Maximum tongue force presented smaller variation than mean tongue force, and the best way to summarize it was by the mean of three measurements. It is recommended that

individuals undergo one or more familiarization sessions with the instrument prior to performance of official measurements.

## ACKNOWLEDGEMENTS

The authors are grateful to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for the funding of this study.

## REFERENCES

1. Furlan RMMM, Motta AR, Valentim AF, Barroso MFS, Costa CG, Las Casas EB. Protrusive tongue strength in people with severely weak tongues. *Int J Speech Lang Pathol*. 2013;15(5):503-10. <http://dx.doi.org/10.3109/17549507.2012.752863>. PMID:23336828.
2. Motta AR, César CC, Bommarito S, Chiari BM. Axial force of the tongue in different age groups. *J Soc Bras Fonoaudiol*. 2011;23(3):201-5. <http://dx.doi.org/10.1590/S2179-64912011000300004>. PMID:22012153.
3. Pittman LJ, Bailey EF. Genioglossus and intrinsic electromyographic activities in impeded and unimpeded protrusion tasks. *J Neurophysiol*. 2009;101(1):276-82. <http://dx.doi.org/10.1152/jn.91065.2008>. PMID:18987117.
4. Adams V, Mathisen B, Baines S, Lazarus C, Callister R. A systematic review and meta-analysis of measurements of tongue and hand strength and endurance using the iowa oral performance instrument (IOP). *Dysphagia*. 2013;28(3):350-69. <http://dx.doi.org/10.1007/s00455-013-9451-3>. PMID:23468283.
5. Ball S, Idel O, Cotton SM, Perry A. Comparison of two methods for measuring tongue pressure during swallowing in people with head and neck cancer. *Dysphagia*. 2006;21(1):28-37. <http://dx.doi.org/10.1007/s00455-005-9008-1>. PMID:16544088.
6. Kieser J, Singh B, Swain M, Ichim I, Waddell JN, Kennedy D, et al. Measuring intraoral pressure adaptation of a dental appliance allows measurement during function. *Dysphagia*. 2008;23(3):237-43. <http://dx.doi.org/10.1007/s00455-007-9126-z>. PMID:18004622.
7. Hewitt A, Hind J, Kays S, Nicosia M, Doyle J, Tompkins W, et al. Standardized instrument for lingual pressure measurement. *Dysphagia*. 2008;23(1):16-25. <http://dx.doi.org/10.1007/s00455-007-9089-0>. PMID:17602265.
8. Hori K, Ono T, Tamine K, Kondo J, Hamanaka S, Maeda Y, et al. Newly developed sensor sheet for measuring tongue pressure during swallowing. *J Prosthodont Res*. 2009;53(1):28-32. <http://dx.doi.org/10.1016/j.jpjor.2008.08.008>. PMID:19318068.
9. Valentim AF, Furlan RMMM, Perilo TVC, Berbert MCB, Motta AR, Las Casas EB. Evaluation of the force applied by the tongue and lip on the maxillary central incisor tooth. *CoDAS*. 2014;26(3):235-40. <http://dx.doi.org/10.1590/2317-1782/201420130077>. PMID:25118921.
10. Furlan RMMM, Valentim AF, Motta AR, Barroso MFS, Costa CG, Las Casas EB. Quantitative methods for assessing tongue force. *Rev CEFAC*. 2012;14(6):1215-25. <http://dx.doi.org/10.1590/S1516-18462012005000099>.
11. Motta AR, Perim JV, Perilo TVC, Las Casas EB, Costa CG, Magalhães FE, et al. Method for the measurement of axial forces produced by the human tongue. *Rev CEFAC*. 2004;6(2):164-9.
12. Barroso MFS, Costa CG, Saffar JME, Las Casas EB, Motta AR, Perilo TVC, et al. Development of a prototype system for the objective measurement of tongue forces. *SBA: Controle & Automação*. 2009;20(2):156-63.
13. Furlan RMMM, Valentim AF, Perilo TVC, Costa CG, Barroso MF, Las Casas EB, et al. Quantitative evaluation of tongue protrusion force. *Int J Orofacial Myology*. 2010;36:33-43. PMID:23362601.
14. Almeida LD, Furlan RMMM, Las Casas EB, Motta AR. Influence of height, weight and body mass index in the axial tongue force. *J Soc Bras Fonoaudiol*. 2012;24(4):381-5. <http://dx.doi.org/10.1590/S2179-64912012000400015>. PMID:23306690.
15. Berbert MCB, Brito VG, Furlan RMMM, Perilo TVC, Valentim AF, Barroso MFS, et al. Maximum protrusive tongue force in healthy young adults. *Int J Orofacial Myology*. 2014;40:56-63. PMID:27295848.
16. Rezende BA, Furlan RMMM, Las Casas EB, Motta AR. Relationship between clinical and instrumental assessment of the tongue in healthy young adults. *CoDAS*. 2015;27(3):260-6. <http://dx.doi.org/10.1590/2317-1782/20152014155>. PMID:26222943.
17. Motta AR, Las Casas EB, César CC, Bommarito S, Chiari BM. Characterization of tongue strength via objective measures. *Rev CEFAC*. 2017;19(1):82-9. <http://dx.doi.org/10.1590/1982-021620171919116>.
18. Furlan RMMM, Las Casas EB, Motta AR. Avaliação quantitativa e reabilitação da força da língua: uma nova ferramenta para a terapia de motricidade orofacial. 1. ed. Saarbrücken: Novas Edições Acadêmicas; 2015.
19. Adams V, Mathisen B, Baines S, Lazarus C, Callister R. Reliability of measurements of tongue and hand strength and endurance using the iowa oral performance instrument with healthy adults. *Dysphagia*. 2014;29(1):83-95. <http://dx.doi.org/10.1007/s00455-013-9486-5>. PMID:24045852.
20. INMETRO: Instituto Nacional de Metrologia, Normalização e Qualidade Industrial. Vocabulário internacional de termos fundamentais e gerais de metrologia. 3. ed. Rio de Janeiro: INMETRO; 2003. Technical Report.
21. Mano T, Katsuno M, Banno H, Suzuki K, Suga N, Hashizume A, et al. Tongue pressure as a novel biomarker of spinal and bulbar muscular atrophy. *Neurology*. 2014;82(3):255-62. <http://dx.doi.org/10.1212/WNL.0000000000000041>. PMID:24353334.
22. Youmans SR, Stierwalt JAG. Measures of tongue function related to normal swallowing. *Dysphagia*. 2006;21(2):102-11. <http://dx.doi.org/10.1007/s00455-006-9013-z>. PMID:16685469.
23. Lazarus C, Logemann JA, Huang CF, Rademaker AW. Effects of two types of tongue strengthening exercises in young normals. *Folia Phoniatr Logop*. 2003;55(4):199-205. <http://dx.doi.org/10.1159/000071019>. PMID:12802092.
24. Clark HM. Specificity of training in the lingual musculature. *J Speech Lang Hear Res*. 2012;55(2):657-67. [http://dx.doi.org/10.1044/1092-4388\(2011/11-0045\)](http://dx.doi.org/10.1044/1092-4388(2011/11-0045)). PMID:22215031.
25. Fernandes A, Marins JCB. Test of hand grip strength: a methodological analysis and normative data in athletes. *Fisioter Mov*. 2011;24(3):567-78. <http://dx.doi.org/10.1590/S0103-51502011000300021>.
26. Mortimore IL, Fiddes P, Stephens S, Douglas NJ. Tongue protrusion force and fatigability in male and female subjects. *Eur Respir J*. 1999;14(1):191-5. <http://dx.doi.org/10.1034/j.1399-3003.1999.14a32.x>. PMID:10489850.
27. Dworkin JP, Aronson AE, Mulder DW. Tongue force in normals and in dysarthric patients with amyotrophic lateral sclerosis. *J Speech Hear Res*. 1980;23(4):828-37. <http://dx.doi.org/10.1044/jshr.2304.828>. PMID:7003262.
28. Youmans SR, Youmans GL, Stierwalt JAG. Differences in tongue strength across age and gender: is there a diminished strength reserve? *Dysphagia*. 2009;24(1):57-65. <http://dx.doi.org/10.1007/s00455-008-9171-2>. PMID:18690406.
29. O'Connor CM, Lowery MM, Doherty LS, McHugh M, O'Muircheartaigh C, Cullen J, et al. Improved surface EMG electrode for measuring genioglossus muscle activity. *Respir Physiol Neurobiol*. 2007;159(1):55-67. <http://dx.doi.org/10.1016/j.resp.2007.05.011>. PMID:17707698.
30. Trawitzki LVV, Borges CGP, Giglio LD, Silva JB. Tongue strength of healthy young adults. *J Oral Rehabil*. 2011;38(7):482-6. <http://dx.doi.org/10.1111/j.1365-2842.2010.02182.x>. PMID:21143618.

## Author contributions

TGA, TMR and RMMM were responsible for the study design, collection and analysis of the data, and writing of the manuscript; ARM and EBLC were the study advisers, responsible for supervising the research project, data analysis, and writing of the manuscript.