DOI: 10.1590/2317-1782/20162015039

Original Article

Artigo Original

Methods of analysis speech rate: a pilot study

Metodologias de análise da velocidade de fala: um estudo piloto

Luanna Maria Oliveira Costa[1]
Vanessa de Oliveira Martins-Reis[1]
Letícia Côrrea Celeste[2]

Keywords

Speech, Language and Hearing Sciences
Speech
Reference Standards
Speech Acoustics
Speech Production Measurement

ABSTRACT

Purpose: To describe the performance of fluent adults in different measures of speech rate. **Methods:** The study included 24 fluent adults, of both genders, speakers of Brazilian Portuguese, who were born and still living in the metropolitan region of Belo Horizonte, state of Minas Gerais, aged between 18 and 59 years. Participants were grouped by age: G1 (18–29 years), G2 (30–39 years), G3 (40–49 years), and G4 (50–59 years). The speech samples were obtained following the methodology of the Speech Fluency Assessment Protocol. In addition to the measures of speech rate proposed by the protocol (speech rate in words and syllables per minute), the rate of speech into phonemes per second and the articulation rate with and without the disfluencies were calculated. We used the nonparametric Friedman test and the Wilcoxon test for multiple comparisons. Groups were compared using the nonparametric Kruskal Wallis. The significance level was of 5%. **Results:** There were significant differences between measures of speech rate involving syllables. The multiple comparisons showed that all the three measures were different. There was no effect of age for the studied measures. These findings corroborate previous studies. **Conclusion:** The inclusion of temporal acoustic measures such as speech rate in phonemes per second and articulation rates with and without disfluencies can be a complementary approach in the evaluation of speech rate.

Descritores

Fonoaudiologia
Fala
Padrões de Referência
Acústica da Fala
Medida da Produção da Fala

RESUMO

Objetivo: Descrever o desempenho de adultos fluentes em diferentes medidas de velocidade de fala. **Métodos:** Participaram do estudo 24 adultos fluentes, de ambos os gêneros, falantes da variante mineira do Português Brasileiro, nascidos e residentes na região metropolitana de Belo Horizonte, Minas Gerais, com faixa etária entre 18 e 59 anos de idade. Os participantes foram agrupados por faixa etária em: G1 (18 a 29 anos), G2 (30 a 39 anos), G3 (40 a 49 anos) e G4 (50 a 59 anos). A coleta da amostra de fala seguiu a metodologia do Protocolo para Avaliação da Fluência da Fala. Além das medidas de velocidade de fala propostas pelo protocolo (taxa de elocução em palavras e sílabas por minuto), foram calculadas: taxa de elocução em fones por segundo e taxa de articulação com e sem as disfluências. Utilizou-se o teste não paramétrico de Friedman e o teste de Wilcoxon para as múltiplas comparações. Os grupos foram comparados por meio do teste não paramétrico de Kruskall Wallis. O nível de significância adotado foi de 5%. **Resultados:** Verificou-se diferença significativa entre as medidas de velocidade de fala que envolvem sílabas e as múltiplas comparações apontaram que as três medidas são diferentes entre si. Não houve efeito da idade para as medidas estudadas. Esses achados corroboram estudos anteriores. **Conclusão:** A inclusão de medidas temporais acústicas, como a taxa de elocução em fones por segundo e taxas de articulação com e sem disfluências, podem ser uma metodologia complementar na avaliação da velocidade de fala.

Correspondence address:
Vanessa de Oliveira Martins-Reis
Departamento de Fonoaudiologia da Faculdade de Medicina da Universidade Federal de Minas Gerais
Avenida Professor Alfredo Balena, 190, sala 251, Santa Efigênia, Belo Horizonte (MG), Brasil, CEP: 30130-100.
E-mail: vomartins@ufmg.br

Received: 02/23/2015

Accepted: 06/02/2015

Study carried out at the Speech-Language Pathology and Audiology Department, School of Medicine, Universidade Federal de Minas Gerais – UFMG – Belo Horizonte (MG), Brazil.
(1) Undergraduate program in Speech-Language Pathology and Audiology, School of Medicine, Universidade Federal de Minas Gerais – UFMG – Belo Horizonte (MG), Brazil.
(2) Universidade de Brasília – UnB – Brasília (DF), Brazil.
Conflict of interests: nothing to declare.

INTRODUCTION

Speech rate is an important measure of fluency of speech, being inversely proportional to the severity of stuttering[1-3]. Moreover, it is a parameter that allows assessment of the motor speech processing[2]. Thus, it is important to characterize the speech rate both in stutterers and in fluent individuals.

Fluency can be defined as the continuous and smooth flow of speech production[2] and is commonly analyzed according to the following parameters: type of breaks (typical disfluencies and stuttering), frequency of breaks (speech discontinuity percentage and percentage of stuttering), and speech rate in words and syllables per minute[1,2,4]. According to the authors of previous studies[1,2,4], the speech rate in words per minute measures the information production rate and, in syllables per minute, the articulatory rate.

However, we found different proposals for the analysis of speech rate in the literature, both with regard to the measurement unit (words per minute, syllables per minute, or phonemes per minute)[1,2] and the exclusion or not of speech disruptions[5], which can often lead the clinician to difficulty in choosing a methodology.

In another theoretical perspective (dynamic systems), there is a temporal unit for analysis other than the syllable, called vowel–vowel (VV) unit. From this perspective, the measures derived from the VV unit also reflect the temporal organization of speech[6]. It is worth highlighting that studies on the fluency of speech, developed by Speech-Language Pathologists, have used this approach before[7]. In this study, we chose to not add it, as it involves another divergent theoretical approach.

The speech rate can be manually measured by stopwatch or by temporal acoustical measures in computerized programs during different speech tasks[8-10]. Structured situations, in which it is possible to control the occurrence of breaks — similar to the repetition of sentences — minimize the influence of the formulation of language in speed and better reflect the performance of the motor mechanisms of speech production[11]. This variety of tasks may interfere with the diagnosis of stuttering, being important to check the best analysis methodology and establish reference values for such.

Some studies[12,13] propose a series of measures related to the temporal organization of speech. Three simple duration measures are proposed as a starting point: time of utterance (total duration of a given statement), time of breaks (total length of silent spaces in the speech), and total articulation time (which is the resulting length of subtracting the time of breaks from the total time of utterance).

On the basis of the three aforementioned duration measures, researchers[13] have proposed temporal variables considering the number of expressed syllables and measures of the total utterance time and total articulation time. The authors proposed that, by dividing the total number of syllables by the total utterance time, one can calculate the speech rate. This gives the listener a global sense of the speech rate. They also proposed that, to obtain the articulation rate, one has to simply divide the total number of syllables by the total articulation time.

These temporal measures and variables were used in further studies of the reading performance of adults[14], of the reading performance of children[15], of the reading performance of children with dyslexia[16], in the speech performance of military commanders[17], in the speech performance of stutterers and fluent individuals[5,18], among others.

When taking measurements of an utterance and its breaks, for example, it is believed that this examination is related to the physical level, as it is limited to assess the time spent in a given space (represented by the speech signal). However, from the moment that the measures relate to the temporal organization of speech, we move to the phonetic level of the analysis. This means that the acoustic parameters taken from the signal and reinterpreted within a linguistic perspective lead us to the phonetic level of the analysis. Inside the level of analysis we find, then, the temporal variables[5].

It should be noted, therefore, the importance of studying the best methodology for assessing speech rate, as it can be used not only to evaluate fluency disorders but also to evaluate communicative effectiveness more broadly. In the literature, there are studies that use different methodologies; but, no study was found comparing the different methodologies. The objective of this study is to describe the performance of fluent adults in different speech rate measures.

METHODS

This is cross-sectional, analytical, observational study with nonprobabilistic sample. The study included 24 adults, fluent speakers of the Minas Gerais variant of Brazilian Portuguese, born and living in the metropolitan region of Belo Horizonte, Minas Gerais, aged 18–59 years, 13 women and 11 men. Participants were grouped according to age: G1 (18–29 years); G2 (30–39 years); G3 (40–49 years); and G4 (50–59 years), with six subjects per age group.

This study was approved by the Research Ethics Committee of Universidade Federal de Minas Gerais (CAAE: 01460612.4.0000.5149). All participants signed an informed consent, which explained the research with its risks and benefits, guaranteeing freedom of participation, refusal or withdrawal, and the confidentiality of personal data.

Participants were asked to participate voluntarily in public places such as parks, schools, universities, among others, located in the metropolitan region of Belo Horizonte, Minas Gerais.

Inclusion criteria were being in the age range determined by the research and having signed the informed consent. Exclusion criteria were adults with personal and/or family complaints of stuttering and other communication disorders; presence of neurological changes and/or psychiatric illness; and who have undergone previous Speech-Language and/or psychiatric therapy. All participants must have been born and lived in Minas Gerais for the past 10 years.

The patient's history was surveyed for the inclusion and exclusion criteria. Participants underwent speech sample collection from a picture as a visual stimulus. They were given the following order: "Please look at this picture and tell me whatever you want about it"[1]. The speech was only interrupted by questions and/or comments, where there was need to encourage the production of discourse to obtain 200 uttered syllables (fluent), required for sample analysis[1].

Speech samples were recorded on a digital mono voice recorder, Sony® ICD-PX333, and transcribed literally, seeking to survey speech disruptions (hesitation, interjection, revision, unfinished word, word repetition, segment repetition, phrase repetition, syllable repetition, sound repetition, prolongation, block, pause, and sound or segment intrusion) and obtaining the speech rate in words and syllables per minute, according to the methodology proposed by the Speech Fluency Protocol[1].

To obtain the speech rate and speech disruption measures, an acoustic analysis was performed using the software Praat, version 5.0.03. To this end, the speech samples recorded on the digital recorder were transferred to a notebook for acoustic analysis.

In the acoustic analysis, the following actions took place: total utterance time (TTEe), break time and unfilled hesitation (TPHnp), pause time (Tp), and the total time of disfluencies (TTdisf).

To calculate the speech rate in words (TxE w/m) and speech rate into syllables per minute (TxE s/m), the methodology proposed by the Speech Fluency Assessment Protocol was used[1], wherein the total number of syllables and words is divided by the total time of utterance and multiplied by 60.

The speech rate in phonemes per second (TxE p/s) refers to the total phonemes of the message uttered, divided by the total utterance time (TTEe). To check the number of phonemes present in the utterance, the phonetic transcription was performed for their numerical count in the 200 syllables uttered.

TTEe refers to the total time spent in the utterance of the sentence (200 syllables uttered). For the calculation, the silence time (pauses and hesitations unfilled) and the time spent in the production of speech disruptions were not accounted for. This measure was performed by acoustic analysis.

The total time of articulation (TTArt) refers to the TTEe minus the time of break and hesitations unfilled (TPHnp) in order to ascertain the exact time used only with articulation. The duration of TPHnp was verified by the acoustic analysis of the utterance.

The articulation rate (TxArt) was calculated with and without the disfluencies. To calculate the articulation rate with disfluencies (TxArtCdisf), the previously selected uttered messages were reexamined and disfluencies produced were computed on the syllable count. Then, all produced syllables (with or without disfluencies) were multiplied by 60, and this result was divided by the time of articulation with disfluency (TArtCdisf). The TArtCdisf refers to TTEe minus the Tp, obtained by the acoustic analysis.

To calculate the articulation rate without dysfluency (TxArtSdisf), the total number of syllables without dysfluency (200 syllables uttered) was multiplied by 60 and divided by the total time of articulation (TTArt). The TArtSdisf refers to the TTEe minus the total time of disfluencies (TTdisf).

Descriptive statistics measures and inferential tests were performed. To compare the speech rate measures involving the syllable count (speech rate in syllables per minute and articulation rates), we used the nonparametric Friedman test and the Wilcoxon test for multiple comparisons. The age groups were compared using the nonparametric Kruskal Wallis test. The significance level was 5%.

RESULTS

Table 1 presents the descriptive statistics of the speech rate measures considered in this study. The Friedman test showed significant difference between the speech rate measures involving syllables (χ^2=45.52; p<0.001) and the multiple comparisons (Wilcoxon test) showed that the three measures are different (Table 2).

It can be seen that there was no effect of age on any of the measures studied (Table 3).

DISCUSSION

This study sought to describe the performance of fluent adults in different speech rate measures. Studies show that the analysis of speech rate has proven valuable in evaluating individuals with typical and altered speech development[1,2,15,16,19-23]. However, there is no consensus on the best methodology for assessing speech rate, which should be associated with the clinician's objective.

In Brazil, the most traditional protocol for assessing speech rate[1] takes into account the utterance rate (words and syllables per minute). This protocol has been used extensively with fluent speakers and individuals with speech and language disorders[1,2,19-23]. The question raised is that the measurement into syllables per minute intends to study the speech rate but does not discount the break time. When the occurrence of breaks is controlled, it minimizes the influence of language formulation in speed, better reflecting the performance of the motor speech-production mechanisms[11].

The average utterance rate found was of 90.25 words per minute and 170.04 syllables per minute. These data corroborate with other studies for Brazilian Portuguese[18,19].

Table 1. Description of the speech rate measures

Speech rate measures	Mean	Median	Standard deviation
Speech rate in words per minute	90.25	89.50	26.348
Speech rate in phonemes per second	6.04	6.00	1.706
Speech rate in syllables per minute	170.04	166.00	44.472
Articulation rate without disfluencies	180.88	175.50	53.249
Articulation rate with disfluencies	189.79	187.50	54.851

Table 2. Result of the multiple comparisons

	TxArtCdisf × TxE s/m	TxArtSdisf × TxE s/m	TxArtCdisf × TxArtSdisf
χ^2	-4.112	-4.289	-4.016
p-value	0.000*	0.000*	0.000*

*Significant values (p<0.05), Wilcoxon test.

Caption: TxArtCdisf = articulation rate with disfluencies; TxE s/m = speech rate into syllables per minute; TxArtSdisf = articulation rate without dysfluency.

Table 3. Effects of age on the speech rate measures

Group (age)	Speech rate (syllables/minute)	Speech rate (words/minute)	Speech rate (phonemes/second)	Speech rate with disfluencies	Speech rate without disfluencies
G1 (18–29 years)					
Mean	196.17	109.33	7.33	205.58	212.50
Median	205.50	112.50	7.50	214.00	221.00
Standard deviation	37.134	25.750	1.366	36.351	37.639
G2 (30–39 years)					
Mean	179.33	94.67	6.33	199.33	209.00
Median	158.50	89.50	6.00	166.00	184.50
Standard deviation	53.549	27.068	1.862	75.437	74.780
G3 (40–49 years)					
Mean	146.33	76.33	5.00	150.83	160.17
Median	139.50	73.00	5.00	143.00	159.00
Standard deviation	24.072	11.827	0.894	25.262	26.309
G4 (50–59 years)					
Mean	158.33	80.67	5.50	167.50	177.50
Median	158.50	80.50	5.50	167.00	181.00
Standard deviation	50.103	29.235	1.871	54.213	61.990
Kruskal Wallis	χ^2=4.624	χ^2=5.610	χ^2=6.361	χ^2=4.715	χ^2=4.317
	p=0.202	p=0.132	p=0.095	p=0.194	p=0.229

The average TxArtSdisf is higher compared with the TxE s/m and with the TxArtCdisf, which is expected, as the disfluencies interfere in this measure.

When the time of dysfluencies is subtracted from the articulation time, the articulation rate aims, in this case, to determine the average duration of each syllable produced and only the effectively produced syllables. This proposal is interesting, because it allows the comparison of the average duration of each syllable of the speech of nonstuttering individuals and the fluent speech of stuttering individuals[5], reflecting the speech rate without influence or with less influence of linguistic processing[11].

The methodology of the inclusion of the length of disfluencies in the articulation time is coherent, as disfluencies are articulatory attempts to produce a particular phoneme. As its name implies, the articulation time includes the entire duration of the moments in which the segments were articulated[5].

The development of motor speech processes extends past the age of 16 years and may be established at around 21 years[23]. In this study, in which the sample consisted exclusively of adults (18–59 years), no effects of age were observed in speech rate measures. These findings corroborate with those of previous studies[19,24], whose results showed that speech rate varies along the stages of life, which may indicate a process of acquisition, development, stabilization, and degeneration. However, it is noteworthy that the previously conducted research[19] observed a decrease in the age group 48–59 years, which differed from the two neighboring groups (38–47 years and 60–69 years), which was not observed in this study. Children speak more slowly compared with adolescence and adulthood. The profile is very close between children and the elderly, initially suggesting maturation of the central nervous system[19].

The findings for speech rate measurements analyzed in relation to age reinforces the findings in the literature on the variability of flow between individuals[5,18,19], even when all are considered fluent, as shown by the high standard deviation values.

This study is a pilot project. By definition, the pilot study is a test on a small scale of procedures, materials, and methods proposed for a given type of research[25]. The importance of conducting a pilot study is in the possibility to test, evaluate, revise, and improve the instruments and research procedures[25,26].

It is believed that this pilot study with fluent speakers has contributed with relevant information. However, it is worth mentioning the importance of further studies to investigate speakers with speech fluency disorders. Only then, it would be possible to establish the best methodology for assessing speech rate.

It would also be interesting that the methodology was applied in other Brazilian regions, allowing the analysis of the influence of “regionality” and “influence of the source language” variables on the speech rate measures, and in a study involving other age groups.

The inclusion of acoustic temporal measures for assessing speech rate proved to be an effective methodology. However, analyzing the cost benefit of these measures in the clinical practice in Speech-Language Pathology and Audiology is of paramount importance, because the analysis proposed by other author[1] is faster than the analyses in this study. Nevertheless, the use of acoustic analysis is suggested to survey the length of the speech samples, as the use of the stopwatch depends on the examiners reaction time and can interfere with the accuracy of the calculation of the speech rate.

CONCLUSION

The results of this study suggest that the inclusion of acoustic temporal measures such as speech rate in phonemes per second and articulation rates with and without disfluencies can be a complementary methodology to evaluate the speech rate.

**LMOC was responsible for collecting and tabulating the data and drafting the manuscript; LCC and VOMR were responsible for the study planning design and overall direction of the execution and preparation of the manuscript.*

REFERENCES

1. Andrade CRF. Fluência. In: Andrade CRF, Befi-Lopes DM, Fernandes FDM, Wertzner HF. ABFW: teste de linguagem infantil nas áreas de fonologia, vocabulário, fluência e pragmática. 2ª ed (revisada, ampliada e atualizada). Barueri: Pró-Fono; 2004. p. 51-82.
2. Andrade CRF, Cervone LM, Sassi FC. Relationship between the stuttering severity index and speech rate. Sao Paulo Med J. 2003;121(2):81-4.
3. Dehqan A, Bakhtiar M, PanahI SS, Ashayeri H. Relationship between stuttering severity in children and their mothers speaking rate. Sao Paulo Med J. 2008;126(1):29-33.
4. Arcuri CF, Osborn E, Schiefer AM, Chiari BM. Taxa de elocução de fala segundo a gravidade da gagueira. Pró-Fono R Atual Cient. 2009;21(1):45-50.
5. Celeste LC, Reis C. Expressão de certeza e dúvida na gagueira: estudo dos aspectos temporais da fala. Rev CEFAC. 2013;15(6):1609-20.
6. Barbosa PA. Análise e modelamento dinâmicos da prosódia do português brasileiro. Rev Estud Ling. 2007;15(2):75-96.
7. Merlo S. Dinâmica temporal de pausas e hesitações na fala semi-espontânea [tese]. Campinas: Universidade Estadual de Campinas; 2012.
8. Oliveira CR, Ortiz KZ, Vieira MM. Disartria: estudo da velocidade de fala. Pró-fono R Atual Cient. 2004;16(1):39-48.
9. Odell KH, Shriberg LD. Prosody-voice characteristics of children and adults with apraxia of speech. Clin Linguist Phon. 2001;15(4):275-307.
10. Walker JF; Archibald LM. Articulation rate in preschool children: a 3-year longitudinal study. Int J Lang Commun Disord. 2006;41(5):541-65.
11. Flipsen P Jr. Articulation rate and speech-sound normalization failure. J Speech Lang Hear Res. 2003;46(3):724-37.
12. Grosjean F. Le rôle joué par trois variables temporelles dans la compréhension orale de l'anglais étudié comme seconde langue, et perception de la vitesse de lecture par des lecteurs et des auditeurs [tese]. Paris: Université de Paris VII; 1972.
13. Grosjean F, Deschamps A. Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de parole et variable composantes: phénomènes d'hesitation. Phonetica. 1975;31(3-4):144-84.
14. Perlman M, Clark N, Falck MJ. Iconic prosody in story reading. Cogn Sci. 2015;39(6):1348-68.
15. Veenendaal NJ, Groen MA, Verhoeven L. The role of speech prosody and text reading prosody in children's reading comprehension. Br J Educ Psychol. 2014;84(4):521-36.
16. Alves LM, Reis CAC, Pinheiro A. Prosody and reading in dyslexic children. Dyslexia. 2015;21(1):35-49.
17. Souza LMC. Prosódia e atitude: o comando militar [dissertação]. Belo Horizonte: Universidade Federal de Minas Gerais; 2007.
18. Andrade CRF, Martins VO. Influencia del sexo y el nivel educativo en la fluidez del habla en personas adultas. Revista de Logopedia Foniatría y Audiología. 2011;31(2):74-81.
19. Martins VO, Andrade CRF. Perfil evolutivo de fala de falantes do português brasileiro. Pró-Fono R Atual Cient. 2008;20(1):7-12.
20. Oliveira CMC, Bernardes APL, Broglio GAF, Capellini SA. Perfil da fluência de indivíduos com taquifemia. Pró-Fono R Atual Cient. 2010;22(4):445-50.
21. Oliveira CMC, Fiorin M, Nogueira PR, Laroza CP. Perfil da fluência: análise comparativa entre gagueira desenvolvimental persistente familial e isolada. Rev CEFAC. 2013;15(6):1627-34.
22. Rossi NF, Souza DH, Moretti-Ferreira D, Giacheti CM. Perfil da fluência da fala na síndrome de Williams-Beuren: estudo preliminar. Pró-Fono R Atual Cient. 2009;21(2):107-12.
23. Andrade CRF, Befi-Lopes DM, Juste FS, Cáceres-Assenço AM, Fortunato-Tavares TM. Aspectos da fluência da fala em crianças com distúrbio específico de linguagem. Audiol Commun Res. 2014;19(3):252-7.
24. Duchin AW, Mysak ED. Disfluency and rate characteristics of young adult, middle-aged, and older males. J Commun Disord. 1987;20(3):245-57.
25. Mackey A, Gass S. Second language research: methodology and design. Mahwah: Lawrence Erlbaum; 2005. p. 43-99.
26. Canhota C. Qual a importância do estudo piloto? In: Silva EE, organizador. Investigação passo a passo: perguntas e respostas para investigação clínica. Lisboa: APMCG; 2008. p. 69-72.