ISSN: 2083-389X
eISSN: 2084-3127

J Hear Sci, 2021; 11(4): 42–47
DOI: 10.17430/JHS.2021.11.4.5

DEVELOPMENT, STANDARDIZATION, AND VALIDATION OF BISYLLABIC PHONEMICALLY BALANCED TAMIL WORD TEST IN QUIET AND NOISE

Contributions:
A Study design/planning
B Data collection/entry
C Data analysis/statistics
D Data interpretation
E Preparation of manuscript
F Literature analysis/search
G Funds collection

Geetha Chinnaraj[A-E,G], Devi Neelamegarajan[A,D-E], Udhayakumar Ravirose[B-F]

Department of Audiology, All India Institute of Speech and Hearing, Mysuru, India

Corresponding author: Geetha Chinnaraj; Department of Audiology, All India Institute of Speech and Hearing, Managangothri, 570006, Mysuru, India; email: geethamysore.cs@gmail.com

Abstract

Background: The present study aimed to develop and standardize a phonemically balanced bisyllabic word test in Tamil for adult listeners.

Material and methods: In total, 1015 bisyllabic Tamil words were collected from different sources; 20 Tamil speakers rated the words for familiarity and 5 experts validated the content. Based on the familiarity rating and content validation, 760 words were shortlisted for phonemic balancing. Then 25 phonemically-balanced lists were prepared with 25 words in each. The prepared lists were presented to 100 normal-hearing listeners at 40 dB SL in quiet, and 30 listeners in noise at −5 dB SNR for the standardization process. The lists were also presented at different sensation levels (SLs) in quiet to 30 listeners to obtain a psychometric function.

Results: The mean speech identification scores (SISs) in adults was 99.8% in quiet. The results revealed no significant difference in SIS across the 25 word lists, indicative of list equivalency. The scores increased as the level increased from 10 to 40 dB SL for all the lists, suggesting homogeneity in difficulty and audibility. However, in noise, only 23 lists were equivalent to each other.

Conclusions: All the test lists can be utilized for testing during audiological evaluation in quiet, and 23 word lists are useful in noise.

Key words: Tamil • speech perception • phonemic balance • word list

OPRACOWANIE, STANDARYZACJA I WALIDACJA DWUSYLABOWEGO, FONEMATYCZNIE ZRÓWNOWAŻONEGO TAMILSKIEGO TESTU SŁOWNEGO W CISZY I W SZUMIE

Streszczenie

Wprowadzenie: Celem badania było opracowanie i standaryzacja fonematycznie zrównoważonego testu słów dwusylabowych w języku tamilskim przeznaczonego dla dorosłych.

Materiał i metody: Z różnych źródeł zebrano w sumie 1015 tamilskich słów dwusylabowych; 20 osób posługujących się tamilskim jako pierwszym językiem oceniło je pod względem znajomości, a 5 ekspertów dokonało walidacji treści. Na podstawie oceny znajomości i walidacji treści sporządzono krótką listę 760 słów do zrównoważenia fonematycznego. Następnie opracowano 25 fonematycznie zrównoważonych list, każda po 25 słów. Przygotowane listy były prezentowane 100 osobom z normalnym słuchem, na poziomie głośności 40 dB SL w ciszy, oraz 30 osobom w szumie przy −5 dB SNR w celu przeprowadzenia standaryzacji. Listy były także prezentowane 30 osobom na różnych poziomach głośności (SL) w ciszy celem uzyskania funkcji psychometrycznej.

Wyniki: Średni wynik identyfikacji mowy (SIS) u dorosłych wyniósł 99,8% w ciszy. Nie zaobserwowano znaczącej różnicy SIS pomiędzy 25 listami słów, co świadczy o ich równoważności. Wynik testu był wyższy, gdy podniesiono poziom głośności z 10 do 40 dB SL dla wszystkich list, co świadczy o ich jednorodności pod względem poziomu trudności i słyszalności. Jednak w szumie tylko 23 listy okazały się równoważne.

Wnioski: Wszystkie opracowane listy słów mogą być stosowane do testów podczas oceny audiologicznej w warunkach ciszy, a 23 listy słów są także użyteczne w szumie.

Słowa kluczowe: • język tamilski • percepcja mowy • zrównoważenie fonematyczne • lista słów

Introduction

Speech audiometry is a collection of behavioral hearing assessment procedures that use speech stimuli, making it an essential test in the audiological assessment battery [1]. One of the tests in speech audiometry is establishing speech identification scores (SISs). The SIS reflects an individual's ability to identify speech at supra-threshold levels and can help in differential diagnosis and selection of amplification devices [2].

The first few materials developed for obtaining SIS were PAL PB-50 word list [3], CID W-22 test, and NU-6 test [4], all of which were in English. Later, speech audiometry materials were developed in different languages such as Russian [5] and Spanish [6]. All the above studies make

use of monosyllabic words. Although monosyllabic words form common test materials, the test can involve bisyllables and trisyllables, depending on the linguistic structure of the language [7].

India is a multicultural and multilingual country with a diverse population. Therefore, it is not feasible to have a single standard test for all the languages of the country. The researchers have recognized the need to develop test materials in several Indian languages. Tamil is a common language in Tamil Nadu and Puducherry, and is the fifth most spoken language in India. The first speech audiometry test in Tamil was developed by Dayalan [8]. The test consists of Tamil phonemically balanced (PB) and spondee word lists for adult listeners. Although this test is most commonly used, the material has not been validated on individuals with hearing impairment. Moreover, Dayalan [8] utilized monosyllables for constructing the word lists. The number of familiar monosyllabic words with consonant endings is low in Tamil. In addition, some of the words in the lists are not in everyday use and hence unfamiliar to listeners (e.g. /pi:r/, /su:l/, /sa:r/, /ja:n/, /ʊa:r/) and thus will affect the accuracy of the test [9]. The word lists also contain some borrowed words (e.g. /bus, /ha:l/, /tin/) and colloquial words (e.g. /pe:i/, /de:i/).

To overcome a few shortcomings of the word list created by Dayalan [8], Kapur [10] developed a list of disyllabic words and Mahima [11] developed PB bisyllabic word lists. The test consists of four lists of phonemically balanced word lists in CVCV combination in Tamil and randomized these lists to make 8 lists. However, randomizing the lists without considering homogeneity of audibility reduces the validity. Apart from the lists mentioned above in Tamil, for adult listeners, a high-frequency speech identification test has also been developed [12].

The available tests in Tamil have only a few sets of word lists. Several applications of speech identification assessment mandate the use of many word lists. For example, during routine hearing aid fitting, to avoid practice effect two or more hearing aid models with different settings need to be assessed using separate word lists in each condition. Similarly, research studies with many variables and test conditions also require multiple word lists. For example, Geetha [13] required 24 different conditions for both ears. Hence, the present study aimed to develop 25 equivalent and homogenous bisyllablic word lists for adult listeners in quiet. Monosyllables are the minimum meaningful unit of a language and are non-redundant, but Tamil is both a vowel and consonant ending language and mono-syllabic words with vowel ending are sparse in Tamil. Therefore this study used bisyllabic words. The study also aimed to validate the developed lists in the presence of noise, as list equivalency is not the same in noise as in quiet [14].

Method

Phase I: Development of phonemically balanced word lists in Tamil

Collection of words

Initially, a pool of 1015 bisyllabic words were collected. These words were collected from various sources including newspapers, magazines, textbooks, novels, storybooks, and a dictionary. Proper nouns or words with cultural differences were not included.

Familiarity rating

The collected pool of words underwent familiarity rating by 20 native Tamil speakers from different socio-economic statuses and education levels. The words were rated using a 5-point rating scale: 5, Most familiar (words are well known and often used in conversation); 4, Moderately familiar; 3, Somewhat familiar; 2, Slightly familiar; and 1, Unknown (words never heard). A total of 930 words with average ratings of 3 to 5 were considered for the next step.

Content validation

Content validation of the selected words was carried out using the scale developed by Shi et al. [15] by five experts working in the field. The experts checked if those words met the criteria regarding absence of emotional, cultural, or religious overlay. At the end of this process, 170 were considered unsuitable and hence excluded. This left a total of 760 words for preparing word lists.

Preparation of phonemically balanced words

Out of the 760 words, 625 words were utilised to construct 25 lists of 25 words each. The completed word lists were phonemically balanced to match the frequency of occurrence of phonemes in Tamil. The phonemic balancing was performed based on the data on frequency of occurrence of phonemes in Tamil [16].

Recording

The constructed word lists were recorded in an acoustically treated room using a personal laptop loaded with Adobe Audition version 3.0 software connected to a condenser microphone. Seven Tamil speakers spoke a sample of 50 words. The recorded sample words were presented to five experts in audiology and five Tamil speakers. They rated the recorded words in terms of naturalness, clarity, pronunciation, and pleasantness on a 4-point rating scale with 0 being poor and 3 the best. A female speaker with the maximum rating was chosen for the final word list recording. The recorded waveforms were digitized with a 16-bit A/D converter at a sampling rate of 44.1 kHz. Each recorded word was normalized to 0 dB using Adobe Audition to give the same RMS power. A calibration tone of 1000 Hz was generated, normalized to 0 dB, and added at the beginning of each word list.

Preparation of word lists in noise

To standardize the word lists in noise, a pilot study was done to find out the SNR at which 50% scores could be obtained. For the pilot study, the recorded word lists were mixed with speech spectrum noise at −7, −5, −3, 0 and +3 dB SNR using a Matlab (version 7.8.0.347) code. The prepared stimuli were presented to 25 individuals (5 listeners for each SNR), and the SNR at which 50% scores were obtained was traced. SNR representing 50% of the scores was obtained at −5 dB SNR level. Hence in the second phase, the word lists

were presented at −5 dB SNR to validate in the presence of noise. The noise was generated by extracting the long-term averaged speech spectrum (LTASS) of all the words using a Matlab code generated by Nike [17].

Phase II: Standardization of word lists

Participants

The recorded word lists were presented in quiet on 100 normal-hearing individuals who were native Tamil speakers aged 18 to 50 years (mean = 29.5; SD = 8.2), after a routine audiological evaluation. A calibrated dual-channel audiometer was utilized for routine audiological evaluation. All the participants had hearing thresholds within 15 dB in both ears and had normal middle ears.

Administration of developed word lists

In total, 100 participants listened to 25 lists in quiet at 40 dB SL (ref: PTA). The words were routed through the calibrated audiometer and delivered through Senheisser HDA-200 headphones. The participants repeated each word and every correct response was given a score of 1 or a score of 0 was given for incorrect responses or failure to repeat the words. To obtain a psychometric function across intensity levels (PI–PB), the word lists were presented at 0, 10, 20, and 30 dB SL to a group of 30 (out of the 100) normal-hearing individuals. The order of presentation of word lists was randomized to avoid an order effect. The word list was presented at 0 dB SL and then at 10 dB SL in the first session, and then testing was done again at 20 and 30 dB SL after a break of 5 days to avoid practice effect. To assess test–retest reliability, 10% of the participants were tested again with all the 25 word lists. Another 30 individuals selected randomly from the first group of 100 listened to the word lists at −5 dB SNR to standardize the test in noise.

Analysis

The data were tabulated and analysed in Statistical Package for the Social Sciences (SPSS) v. 21. A Shapiro–Wilks test of normality revealed that the data were not normally distributed. Hence, non-parametric statistics (Friedman test) was carried out. Cronbach's alpha measures were used to assess the homogeneity and the intra-class correlation coefficient. Cronbach's alpha measures were also used to assess test–retest reliability.

Results

Standardization of the developed word lists and assessing consistency

For standardization of the 25 developed word lists, the SIS was obtained at 40 dB SL (re: hearing threshold) on 100 normal-hearing individuals. The number of correctly identified words (hereafter referred to as SIS) for each list was calculated. The mean and SD of the SIS at 40 dB SL are given in Table 1.

Table 1. Mean and SD of speech identification scores for 25 lists at 40 dB SL obtained on individuals with normal hearing sensitivity

Lists	Mean	SD
L1	24.96	0.19
L2	24.93	0.25
L3	24.96	0.19
L4	24.95	0.19
L5	24.98	0.14
L6	24.98	0.14
L7	24.95	0.26
L8	24.97	0.17
L9	24.98	0.14
L10	24.94	0.23
L11	24.97	0.17
L12	24.97	0.17
L13	24.99	0.10
L14	24.95	0.26
L15	24.98	0.14
L16	24.96	0.19
L17	24.96	0.19
L18	24.94	0.23
L19	24.96	0.19
L20	24.94	0.23
L21	24.98	0.14
L22	24.94	0.23
L23	24.99	0.10
L24	24.99	0.10
L25	24.99	0.10

Note: Maximum possible score = 25; L = List

The grand mean average across the lists was 99.77%. It can be seen in Table 1 that the mean number of correctly repeated words ranged between 24.93 to 24.99. A Friedman test was carried out to see whether there was a significant difference in the SIS across 25 words lists. The results revealed no significant difference ($\chi^2 = 28.527$, $p > 0.05$) across the 25 word lists, indicating list equivalency.

Then 10 out of 100 individuals were tested again with all the word lists to assess the test–retest reliability. The data were statistically compared using Cronbach's alpha to check reliability. Cronbach's alpha of 0.86 was obtained, which denotes that the developed lists have high test–retest reliability.

Assessment of PI–PB function

A psychometric function was drawn using a PI–PB test to assess homogeneity in terms of difficulty level across different lists on 30 individuals at 0, 10, 20, 30, and 40 dB SL (Figure 1). Figure 1 shows a gradual increase in SIS with increase in intensity, with a poor score of 0 to 4% at 0 dB SL and 100% scores above 30 dB SL. The same trend was followed by all the lists.

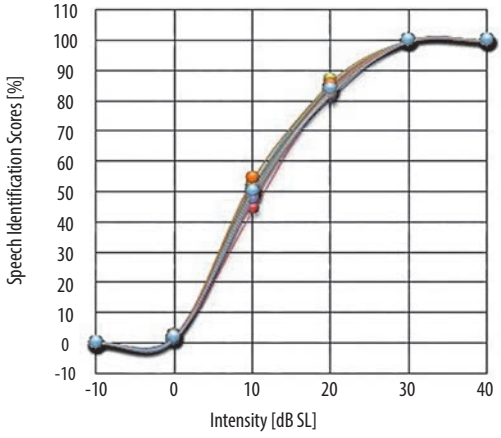

Figure 1. Psychometric functions for the 25 word lists

Homogeneity of audibility was also measured using Cronbach's alpha test. McGraw and Wong [18] reported that the measure of Cronbachs alpha value using intra-class coefficient at the 50% and 100% performance level provide information on the homogeneity of the word lists in terms of audibility. In the present study, 50% score was obtained at 10 dB SL and a ceiling was observed at 40 dB SL. Hence, the SIS at these two levels was employed for homogeneity assessment using Cronbach's alpha. The results revealed that all the word lists were in strong agreement in terms of audibility at 10 dB SL ($\alpha = 0.984$) and at 40 dB SL ($\alpha = 0.987$).

Validation of the word test in noise

The mean and standard deviation of SIS obtained in noise is given in Figure 2. A Friedman test was done to compare SIS across different word lists. Results revealed a significant difference in performance across the lists [$\chi^2(24) = 39.338$, $p < 0.05$].

The results of the pair-wise comparison using Wilcoxon signed-rank test revealed that lists 14 and 25 had significantly different scores from many other lists (given in Table 2) whereas list 20 was significantly different from one list. Lists 5 and 19 were significantly different from two other lists, and list 8 was significantly different from three other lists. However, the analysis of the effect size (η_p^2) revealed that η_p^2 varied between 0.36 to 0.43 for lists 5, 8, 19, and 20, which represents only a small effect size. Hence, lists 5, 8, and 19 can be considered equivalent.

Discussion

The present study aimed to develop 25 PB word lists in Tamil and standardize them in normal-hearing individuals. Results showed that the overall mean combining all the 25 word lists was 99.8%. These scores are comparable with that obtained for word lists developed by Manjula et al. [19] and Mahima [11], who reported a mean score of 98% for Kannada PB word lists and PB words lists developed in Tamil, respectively. Ullrich and Grimm [20] administered NU-6 word lists in normal-hearing individuals and reported a score of 99.7%. Beattic et al. [21] obtained a score of approximately 95% at 32 dB SL using CID W-22 and NU-6 test materials. The reason for such high scores at MCL in normal-hearing individuals is the intact auditory system and attention. Hence, above the most comfortable listening level, almost all normal-hearing individuals achieve a 100% score [22].

It is very important that consistency assessment is done, as it guarantees that the SIS is comparable irrespective of the list used. In the current study, the mean SIS across the lists was statistically comparable and hence they can be used interchangeably across different test conditions.

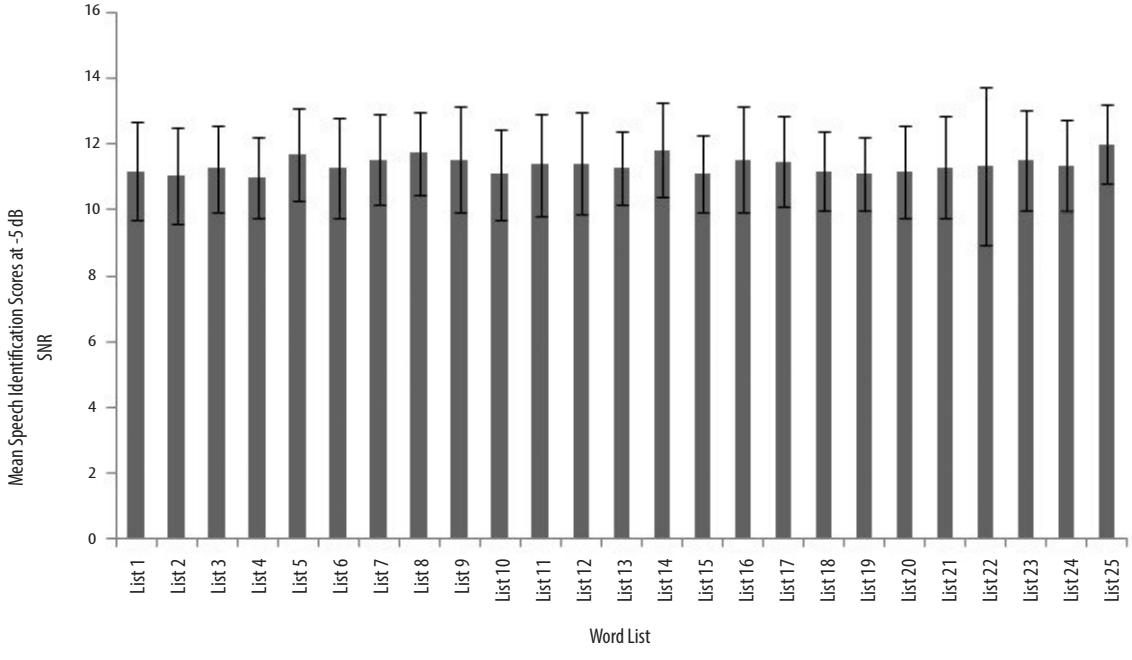

Figure 2. Mean and standard deviation of Speech Identification Scores obtained in noise at –5 dB SNR

Table 2. Bonferroni pair-wise comparison across 25 PB word lists in noise

Lists A	Lists B	*Z*-value	Value of significance (*p*)	effect size (η_p^2)
L14	L1	−2.122	0.034*	0.38
	L2	−3.192	0.001**	0.57
	L3	−2.477	0.013*	0.45
	L4	−3.237	0.001**	0.58
	L10	−2.237	0.025*	0.40
	L13	−2.399	0.016*	0.43
	L15	−2.328	0.020*	0.42
	L18	−2.604	0.009**	0.47
	L19	−2.755	0.006**	0.50
	L20	−2.355	0.019*	0.42
	L21	−2.336	0.019*	0.42
L5	L2	−1.986	0.047*	0.36
	L4	−2.179	0.029*	0.39
L8	L2	−2.413	0.016*	0.43
	L4	−2.381	0.017*	0.43
	L19	−2.372	0.018*	0.43
L19	L7	−2.015	0.044*	0.36
	L22	−2.077	0.038*	0.37
L20	L16	−2.128	0.033*	0.38
L25	L1	−2.518	0.012*	0.45
	L2	−3.054	0.002**	0.55
	L3	−2.647	0.008**	0.48
	L4	−3.034	0.002**	0.55
	L6	−2.524	0.012*	0.45
	L10	−2.926	0.003**	0.53
	L11	−2.352	0.019*	0.42
	L12	−1.998	0.046*	0.36
	L13	−2.254	0.024*	0.41
	L15	−3.283	0.001**	0.59
	L17	−2.231	0.026*	0.40
	L18	−2.636	0.008**	0.47
	L19	−3.010	0.003**	0.54
	L20	−2.775	0.006**	0.50
	L21	−2.263	0.024*	0.41
	L24	−2.149	0.032*	0.39

Note: *$p < 0.05$; **$p < 0.01$; only the significant pair-wise comparisons are shown; L = List

A psychometric equivalency of the lists was also assessed across different intensities. The sigmoid curves were coincident for all the lists, suggesting homogeneity of the word lists. At 10 dB SL, the score was 50.56%. With increase in intensity, the scores improved and reached a ceiling above 30 dB SL. Manjula et al. [19] and Mahima [11] also obtained approximately 50% scores at 10 dB SL and observed a plateau after 30 dB SL.

Even though the PI–PB function curve acts as a measure to assess the list equivalency of a developed test list, assessment of homogeneity in terms of audibility will strengthen the consistency of the developed word lists [7]. Hence, in the present study the homogeneity of audibility was ensured statistically using Cronbach's alpha measures. Alisaputri [7] measured SIS at two dial settings (15 dB and 40 dB SL) in the Malay language. The homogeneity of audibility of Malay PB word lists was assessed using Cronbach's alpha, and the results revealed a high alpha value of 0.81. The results of the study current study are in agreement with that of Alisaputri [7].

Many audiological applications, including hearing aid fitting and research studies, require the SIS be done in the presence of noise. Therefore, it was essential to validate the developed word test in the presence of noise. The SIS using 25 word lists in noise was obtained at −5 dB SNR on 30 normal-hearing individuals. The results revealed no significant differences for 19 lists. Although lists 5, 8, 19, and 20 were significantly different from a few other lists, there was a small effect size, and hence these lists were retained. However, lists 14 and 25 were found to have significantly different scores from many other lists, hence, these two lists cannot be used while testing in the presence of noise. After removing lists 14 and 20, the mean SIS at −5 dB SNR was 11.33 (45.3%). Manjula [19] also obtained similar results for their word lists, that is, 46.0% at −3 dB SNR. In comparison, the 50% mark is at 1 dB SNR for the lists CID W-22, NU-6, and W-1 spondaic words [23]. The performance in the current study is slightly better than that obtained by Wilson et al. [23]. The reason for this difference could be because the former study included monosyllables, which are more difficult to perceive than bisyllables.

Conclusions

The 25 PB word lists developed in Tamil were found to be useful to assess SIS in quiet, and the lists are interchangeable during the course of testing, as there is homogeneity between the lists in terms of difficulty and audibility. There are 23 word lists (omitting 14 and 25) that can be used for speech audiometry tests in the presence of noise.

References

1. Martin FN, Clark JG. Introduction to Audiology. 13th ed. Pearson: 2019.
2. Lawson GD, Peterson ME. Speech Audiometry. San Diego: Plural Publishing; 2011.
3. Egan JP. Articulation testing methods. Laryngoscope, 1948; 58: 955–91. https://doi.org/0.1288/00005537-194809000-00002
4. Tillman TW, Carhart R. An Expanded Test for Speech Discrimination Utilizing CNC Monosyllabic Words: Northwestern University Auditory Test No. 6. Northwestern Univ, Evanston IL Auditory Research Lab; 1966.
5. Aleksandrovsky IV, McCullough JK, Wilson RH. Development of suprathreshold word recognition test for Russian-speaking patients. J Am Acad Audiol, 1998; 9: 417–25.

6. Christensen LK. Performance intensity functions for digitally recorded Spanish speech audiometry. PhD Thesis. Brigham Young University. Department of Educational Psychology, 1995.
7. Alisaputri M. Development of Bisyllabic Speech Audiometry Word Lists for Adult Malay Speakers. Ph.D Thesis. De Montfort University, 2016.
8. Samuel JD. Development and Standardization of Phonetically Balanced Test Materials in Tamil Language. Unpublished Dissertation. AIISH, 1976.
9. Zubick HH, Irizarry LM, Rosen L, Feudo P, Kelly JH, Strome M. Development of speech-audiometric materials for native Spanish-speaking adults. Audiology, 1983; 22: 88–102. https://doi.org/10.3109/00206098309072772
10. Kapur YP. Needs of the Speech and Hearing Handicapped in India: A Study. Project Report of an Investigation of the Audiological and Rehabilitation Needs of Persons with Speech and Hearing Disorders (1966–1970). Christian Medical College and Hospital; 1971.
11. Mahima S, Muthuselvi T. Development of phonemically balanced word list in Tamil for speech audiometry and evaluation of its effectiveness in adults. Unpublished Dissertation. Sri Ramachandra University, 2017.
12. Sinthiya K, Sandeep M. High frequency speech identification test in Tamil. Student Research: Article Based on Dissertation. AIISH, 2009.
13. Geetha C. Optimization of compression parameters in hearing aids using aided audibility index. Doctoral Thesis. University of Mysore, 2016.
14. Chermak GD, Wagner DP, Bended RB. Interlist equivalence of the Word Intelligibility by Picture Identification Test administered in broad-band noise. Audiology, 1988; 27: 324–33. https://doi.org/10.3109/00206098809081603
15. Shi J, Mo X, Sun Z. Content validity index in scale development. Zhong Nan Da Xue Xue Bao Yi Xue Ban, 2012; 37: 152–5. https://doi.org/10.3969/j.issn.1672-7347.2012.02.007
16. Thennarasu S. A statistical study of Tamil corpus. South Asian Lang Rev, 2008; XVIII.
17. Nike G. Speech in noise mixing, signal to noise ratio. Matlab Cent File Exch 2021. https://www.mathworks.com/matlabcentral/fileexchange/37842-speech-in-noise-mixing-signal-to-noise-ratio
18. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods, 1996; 1: 30–46. https://doi.org/10.1037/1082-989X.1.1.30
19. Manjula P, Antony J, Kumar KSS, Geetha C. Development of phonemically balanced word lists for adults in the Kannada language. J Hear Sci, 2015; 5: 22–30.
20. Ullrich K, Grimm D. Most comfortable listening level presentation versus maximum discrimination for word discrimination material. Audiology, 1976; 15: 338–47. https://doi.org/10.3109/00206097609071793
21. Beattie RC, Edgerton BJ, Svihovec DV. A comparison of the Auditec of St. Louis cassette recordings of NU-6 and CID W-22 on a normal-hearing population. J Speech Hear Disord, 1977; 42: 60–4. https://doi.org/10.1044/jshd.4201.60
22. Davis H, Silverman SR. Hearing and Deafness. New York: Holt, Rinehart, and Winston; 1960.lson RH, McArdle R, Roberts H. A comparison of recognition performances in speech-spectrum noise by listeners with normal hearing on PB-50, CID W-22, NU-6, W-1 spondaic words, and monosyllabic digits spoken by the same speaker. J Am Acad Audiol, 2008; 19: 496–506. https://doi.org/10.3766/jaaa.19.6.5