Development and Validation of Tulu Sentence Lists to Test Speech Recognition Threshold in Noise

Shreepriya Bhat, Mohan Kumar Kalaiah, Usha Shastri

Department of Audiology and Speech Language Pathology, Kasturba Medical College, Mangalore, Manipal Academy of Higher Education, Manipal, Karnataka, India

Abstract

Introduction: We aimed to develop sentence lists in Tulu language to measure speech recognition thresholds (SRTs) in noise and verify the equivalence of the developed lists in noise among individuals with normal hearing. **Methods:** This cross-sectional study was conducted in three phases; first, the developmental phase, collection, and development of the sentence material in the Tulu language were carried out. Sentences rated highly familiar/familiar and highly natural/natural by 80% of the raters were considered for the next phase. In the second phase, 22 sentence lists were prepared, with each list containing ten sentences of equal difficulty level in noise. The difficulty level was matched mainly based on the similarity (±1 standard deviation [SD]) of signal-to-noise ratio 50 (SNR-50), the slope of the sentences, the number of content words, and phonemes. The third phase evaluated the equivalency of developed sentence lists. Sixty-two native Tulu speakers with normal hearing sensitivity, aged between 18 and 39 years, participated. Of these 62 participants, twelve were recruited for the first phase, 20 individuals participated in the second phase, and 30 participated in the third phase. **Results:** Thirteen sentence lists were equivalent and were a reliable measure of SNR-50 in native speakers of Tulu who have hearing sensitivity within normal limits. Using the sentence scoring method, the average SNR-50 of these sentence lists was − 4.19 dB with a SD of 0.21 dB. The developed sentences also had less within-subject variability than between-subject variability. **Conclusions:** These 13 equivalent sentence lists provide a valid and reliable tool to measure the SRT in noise in native Tulu speakers with hearing sensitivity within normal limits.

Keywords: Sentence scoring, signal-to-noise ratio-50, speech recognition threshold, Tulu language

Date of Submission: 04-09-2021 **Date of Revision:** 12-10-2021 **Date of Acceptance:** 14-11-2021 **Date of Web Publication:** 24-12-2021

Introduction

Assessment of speech perception is a crucial part of the audiological test battery. Information about an individual's sensitivity to speech stimuli and the ability to understand speech is provided by speech audiometry. Speech detection/awareness threshold, speech recognition threshold (SRT), and speech discrimination/identification score are the commonly used test procedures in a standard speech audiometric procedure. Among these, SRT is the most commonly used procedure in speech audiometry. Optimal test material must be selected to measure SRT for the effective clinical outcome because SRT involves understanding speech stimuli.[1]

Among various materials used to assess the speech perception abilities of an individual, sentences provide information on understanding ability of speech stimuli and provide useful information in assessing communication difficulties experienced by listeners with hearing impairment.[1] Sentences are advantageous as they provide information regarding the contextual characteristics of conversational speech. Furthermore, sentence materials are more reliable and efficient for SRT measurements because of their steeper psychometric functions.[2,3] However, using sentences as stimuli also has disadvantages. To understand sentences, the presence of many contextual cues acts as extrinsic redundancies. Hence, sentences are easy to be memorized. Therefore, in tasks that require repeated measurements of SRT, utilizing sentences is highly difficult. Although disadvantages are present, the advantages outweigh the disadvantages as the sentence

Access this article online

Quick Response Code:

Website: www.jisha.org

DOI: 10.4103/jisha.jisha_22_21

Address for correspondence: Dr. Usha Shastri,
Department of Audiology and Speech-Language Pathology, Kasturba Medical College, Mangalore - 575 001, Karnataka, India.
E-mail: shastri.usha@gmail.com

How to cite this article: Bhat S, Kalaiah MK, Shastri U. Development and validation of tulu sentence lists to test speech recognition threshold in noise. J Indian Speech, Hear Assoc 2021;35:50-7.

material provides a comprehensive view of an individual's speech perception compared to any other material. When sentence material is used, generally, testing is done in background noise as it represents listening difficulties in everyday life situations.[4,5]

Studies have used a variety of methods in developing sentence materials for speech audiometry. First, many sentences can be used to develop different sentence lists,[1,2,6,7] or matrix sentences could be constructed.[8-10] Second, the types of background noise used are different across studies. Investigators have used the long-term average spectrum (LTAS) of the recorded speech material[11,12] and multi-talker babble.[13] Among different types of noise that can be used in the background, the LTAS of the recorded speech material is said to be more accurate in measuring sentence recognition thresholds while using the sentence scoring method[1] the SRT in noise. Some researchers have used an adaptive procedure to vary the signal-to-noise ratio (SNR) depending on the response of the participant.[1,2,7] Others have estimated scores for each sentence in a list at different SNRs.[14,15] The psychometric function of sentence recognition score measured at different SNRs is generally used to form equally intelligible sentence lists. The information from the psychometric function can be utilized in two different ways to form the equivalent lists. One such method is to use the slope of psychometric function to equate the intelligibility of sentences.[12,16] In this method, the difference in an individual sentence's percent intelligibility score from the overall mean will be used to rescale the intensity level of the sentence. Another method is to eliminate the sentences whose slope of the psychometric function and SRTs are significantly deviating from that of other sentences.[1,2,17] We used the later procedure in this study.

The accuracy and validity of speech audiometry can be enhanced by using standardized speech stimuli in the native language of an individual.[18,19] Sentence materials have been developed for speech audiometry in English[12,15] and other foreign languages, for example, German,[2] Canadian French,[20] French,[21] Swedish,[22] Mandarian,[23] Polish,[10] Afrikaans,[24] and Cantonese.[25] India has vast diversity with a multilingual and multicultural population with nearly 29 independent languages. Hence, many researchers have developed materials for speech audiometry and perception tests in Indian languages. For example, sentence lists have been developed in Kannada,[6] Hindi,[7] and Telugu[1] to measure SNR-50. Tulu language, one of the oldest Dravidian languages with 1,846,427 native Tulu speakers in India,[26] is spoken in Dakshina Karnataka and the northern part of Kerala.[27,28] Many regional and social differences lead to five regional dialects and four social dialects of Tulu.[29] It has few materials for assessing speech perception. Samuel[30] developed speech audiometry stimuli in the Tulu language, which are word-level. Sentences lists have not been developed in the Tulu language.

The present study aimed to develop equivalent sentence lists in the Tulu language to measure SRTs in noise. We also aimed to verify the equivalence of developed sentence lists in noise among individuals with normal hearing. As mentioned above, there are various regional and social dialects in Tulu, and each of the dialects has subtle differences among them. Among all the dialects, the common dialect of Tulu is used by the majority of Tulu speakers. In addition, media, magazines, textbooks, and movies of Tulu use common dialect. Hence, people from other regional and social dialects of Tulu are familiar with the common dialect of Tulu. Thus, the sentence material developed in this study used the common dialect of Tulu.

Materials and Methods

Participants

In the study, 62 native Tulu speakers (23 males and 39 females) with normal hearing sensitivity (average hearing threshold ≤25 dB HL in both ears), irrespective of their dialects, aged between 18 and 39 years, participated. Each participant was subjected to pure tone audiometry by utilizing the modified Hughson-Westlake method[31] across octave frequencies between 250 Hz and 8000 Hz for air conduction measurement and 250 Hz 4000 Hz for bone conduction measurement. The pure-tone average in both ears of all the participants was ≤25 dB HL with "A" type tympanogram, and acoustic reflex thresholds for pure tones at octave frequencies from 500 Hz to 4000 Hz were within normal limits. None of the participants reported any complaints/history of cognitive/otological issues or difficulty understanding speech. These participants were distributed among the three phases of the study. All the participants were informed in detail about the procedure, purpose, and need for the study, and written informed consent was taken. They participated voluntarily in the study and were not paid. Ethical clearance from the institutional ethics committee was obtained before commencing the study.

Procedure

The research was done in three phases. The first phase was the developmental phase, wherein the collection of sentences and development of the sentence material in the Tulu language was carried out; in the second phase, sentence lists of equal difficulty level in noise were prepared. The third phase evaluated the equivalency of the developed sentence lists.

Phase 1: Developmental phase

Selection of sentences

A total of 500 sentences were accumulated from various sources like consulting linguists with good knowledge of the Tulu language, few natively available Tulu textbooks, magazines, general conversations, and picture descriptions. The sentence complexity and structure of the sentences were kept as simple and similar as possible. Developed sentences were made semantically neutral by avoiding topics that included politics, gender equality, caste, and war. Ambiguous sentences, sentences including proverbs, questions, exclamation marks were not included in the list. Each sentence had seven to thirteen syllables and three to five words. Developed sentences were syntactically correct and semantically neutral. No duplicate/redundant sentences were selected. Based on the criteria mentioned above, 488 sentences with two to five keywords were selected.

Selected sentences were typed and printed in Kannada script (Tulu script is not commonly used, and Kannada script is generally used in print communication) and were presented to 12 native Tulu speakers irrespective of dialects. They assessed the naturalness and familiarity of every sentence using a subjective rating scale. Familiarity was assessed based on familiarity and occurrence of the sentence in daily conversation on a three-point rating scale (very familiar, familiar, or unfamiliar).[32] Naturalness was rated on a 5-point rating scale, where a score of 5 indicates highly natural, and a score of 1 indicates totally artificial and not heard.[1,6,7] The sentences rated as familiar and highly familiar for familiarity and natural and highly natural for naturalness by at least 80% of the participants were selected for the second phase. Four hundred and forty-one sentences were obtained based on the rating measures. The features of developed sentences, such as sentences representing daily conversation, free from proper nouns, questions, proverbs, and exclamations, warranted that the sentences received a high rating of naturalness and familiarity and only a few exclusions were required.

Recording and editing of the sentences

The sentences were digitally recorded in a sound-treated room using Audacity® version 2.0.5 software at a sampling rate of 48 k Hz with 64-bit resolution. A windshield attached to a powerful directional microphone was placed 10–15 cm from the speaker's mouth. Three adult female speakers in the age range of 21 and 25 years were considered as potential speakers. These three speakers recorded ten sentences each. Their speech sample was given to five audiologists to select a speaker based on their assessment of the voice quality, sentence intonation, rate of speech, stress, intelligibility, and pronunciation. A 21-year-old adult female who was a native Tulu speaker from common dialect was selected as speaker based on the assessment from all the listeners. Each sentence was read out thrice and was recorded and stored in separate wave files in.wav format. One of the three audio samples for each sentence with the best perceptual intelligibility and stable suprasegmentally was chosen as the final test material. These individual wave files were normalized by adjusting the root mean square (RMS) level with respect to maximum digital output. This editing was performed using Adobe Audition, V3 (Adobe Systems Incorporated, San Hose, CA) software.

Generation of noise

Background noise was generated using MATLAB function to equate the LTAS of the 441 sentences. The sentences were concatenated in random order, and a fast Fourier transform (FFT) algorithm was performed. This provides the exact spectral weightage of the sentences, which were multiplied with random phase, and an inverse FFT (iFFT) was performed. Finally, taking the real values of iFFT resulted in the noise, which had a frequency spectrum similar to the LTAS of 441 sentences. Thus, the generated noise had an RMS level that was matched to that of the sentences. Figure 1 shows the LTAS of selected 441 sentences and speech spectrum-shaped noise.

Phase 2: Selection of sentences with equal intelligibility in background noise

Participants and procedure

The recorded sentences were mixed with noise and were presented at four different SNR levels (−2, −4, −6, and −8 dB). These four SNR values were chosen based on the results of previous literature, which reports the mean SRT to be approximately −4 dB SNR[1,7,12,16,20,22] to attain a 50% sentence recognition score. A laptop with MATLAB software version 7.10.0 (MathWorks, USA), Sennheiser HD 280 Pro headphones, and FiiO E17 Alpen Portable Headphone Amplifier were used to present the stimuli. The noise level was kept constant at 60 dB SPL, and the intensity of the sentences was varied to achieve desired SNR. The sentences were presented monaurally to either the participants' right ear or left ear, and the ear was selected randomly.

All the 441 sentences were presented to 20 native Tulu-speaking participants such that one participant was tested only at one SNR, and a total of five participants were tested at each SNR. The participants reported the sentences as fully as possible. Scoring for word recognition was based on total content words repeated correctly. Participants speculated the words if not sure. The percentage of correct identification score was calculated at each SNR for every sentence based on total correctly repeated words, and the identification scores were averaged across the participants. Finally, an average identification score across all 441 sentences was determined at every SNR. Using a logistic function, sentence-specific psychometric curves were derived for each sentence. SNR-50 and slope of identification performance were obtained for all the sentences using the equation below,[2]

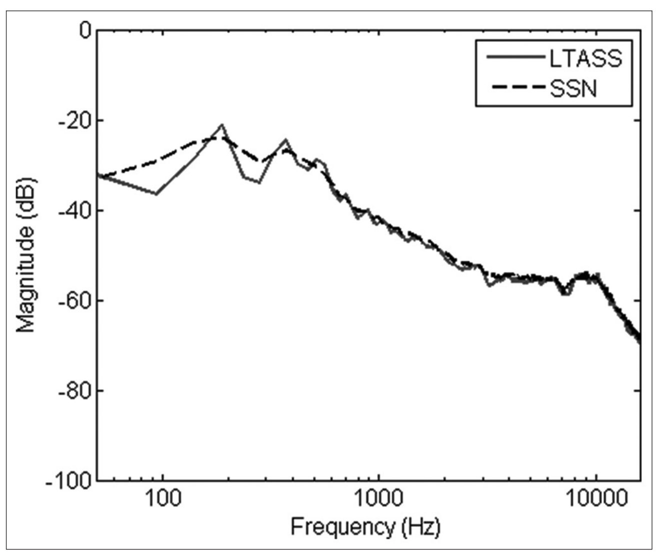

Figure 1: Long-term average spectrum of selected 441 sentences and spectrum of speech spectrum shaped noise

$$P(SNR) = 100 \div 1 + e^{-(SNR-SNR50)/S}$$

Here P (SNR) represents the sentence recognition score in percentage at SNR, SNR indicates the SNR used for sentence identification, SNR50 represents the SNR at which 50% intelligibility was obtained, and 'S' represents the spread of the psychometric curve for every sentence. 'm' represents the slope of the psychometric function, and the spread of the psychometric curve 'S' is inversely proportional to "m" (S = 25/m). Out of 441 sentences, a total of 220 sentences had slope and SNR-50 within ±1 SD on estimated psychometric curves and were considered for the next phase.

Formulation of equally intelligible sentence lists

Twenty-two lists, each having ten sentences, were formulated from the selected 220 sentences. The aim was to obtain several sentence lists with similar SNR-50, slope, number of content words, and phonemes. The lists were formed using a similar procedure described by Tanniru *et al*.[1] and Kollmeier and Wesselkamp.[2] Among all the desired parameters to be matched, SNR-50 and the slope of the sentences were given more importance. The average SNR-50 of all the sentences was considered as the desired SNR-50 for each sentence list. Then, ten sentences were randomly selected, and the average SNR-50 was computed. If the difference between the average of each parameter of the formed sentence list was within ±1 SD of the desired values for all the parameters, these ten sentences were considered one list. For example, if the desired SNR-50 is -3.56 dB and the standard deviation (SD) is 0.32, the SNR-50 obtained from the ten sentences in a list is −3.74 dB (which is within −3.56 ± 0.32), then these ten sentences are considered to be one list. These 10 sentences were eventually removed from the group of 220 sentences. This optimization procedure was repeated using the MATLAB software until 22 lists were formed.

Phase 3: Verification of equivalency of sentence lists

Participants

Thirty native Tulu speakers participated in this phase. All the lists were presented to every participant in random order, and SNR-50 for each list was obtained using the adaptive method. When testing each list, the first sentence was presented at -8 dB SNR. If the first sentence was not repeated correctly, the same sentence was repeated, wherein the signal level was made better by 2 dB. This was continued until the correct response was obtained for the first sentence. Following this, every correct response resulted in reducing the SNR by 2 dB, and every incorrect response resulted in increasing SNR by 2 dB. That is, the one up, one down procedure was used for sentences, keeping noise level constant. Thus, the presentation level was varied in 2 dB steps depending on the listener's response to obtain SNR-50. APEX 3 program developed at ExpORL[33] was customized to do the testing in the adaptive procedure. To score a sentence as correct, the participant had to repeat the whole sentence correctly. SNR-50 was calculated as the average of the last four reversals. The participants were tested monaurally by randomly testing the right ear for 50% of the participants and the left ear for the other 50%.

Statistical analysis

The data for each phase were tabulated, and descriptive statistics were done. Inferential statistics was done for phase 3 data using SPSS Statistics for Windows, Version 17.0. Chicago: SPSS Inc. That is, repeated-measures ANOVA for lists was done to check for equivalence of developed lists.

Results

Selection of sentences with equal intelligibility in background noise

The average and SD of the percentage of correctly identified words across all the 441 sentences at four different SNRs are given in Table 1. The mean slope of the sentences was 23.34% per dB SNR (SD ± 0.24 dB). The mean SNR-50 was −3.56 dB (SD ± 0.32 dB). Among 441 sentences, 220 had slope and SNR-50 within ±1 SD on estimated psychometric curves; hence, they were chosen to prepare sentence lists. Table 1 also provides the percentage correct scores obtained for these 220 selected sentences at four different SNRs.

Verification of equivalency of sentence lists

Average SNR-50 and SD of developed 22 lists obtained across all the participants are given in Table 2. The mean of SNR-50 across 22 lists was −4.28 dB. From Table 2 it can be observed that among 22 lists, lists 1, 7, 11, 12, 14, 15, 18, 20, and 22 had a mean SNR-50 greater or lesser than 0.5 dB of −4.28 dB. This might indicate that the SNR-50 obtained using these lists could differ from the SNR-50 obtained from other lists. Hence, these nine lists were removed, and the remaining 13 lists (sentence lists numbered 2, 3, 4, 5, 6, 8, 9, 10, 13, 16, 17, 19, and 21 were subjected to repeated-measures ANOVA with the list as a within-subject factor. The results showed no significant difference among 13 sentence lists (F [12,348] = 1.15, P = 0.32). Hence, these 13 lists can be considered equivalent and as final test material. The final developed lists were renumbered as lists 1 through 13 instead of the original list numbers. List 1 is given in the Appendix. The average number of words, the average

Table 1: The mean and standard deviation of percent correct word identification score for sentences at four different signal-to-noise ratio across 441 sentences and selected 220 sentences

Percent correct scores	Mean (%)±SD			
	−2 dB SNR	−4 dB SNR	−6 dB SNR	−8 dB SNR
Of 441 sentences	84.53±11.88	37.45±14.41	9.19±8.71	2.51±5.08
Of selected 220 sentences	84.74±7.66	36.87±8.87	8.05±6.66	1.21±3.13

SD: Standard deviation, SNR: Signal-to-noise ratio number of content words, the average number of syllables, and the average number of phonemes per sentence in 13 lists are tabulated in Table 3. The mean SNR-50 obtained from 13 lists was −4.19 dB (SD ± 0.21 dB). The average slope of 13 lists was 21.45% per dB SNR (SD ± 0.66 dB) with a range of 14%–27% per dB SNR.

To ascertain whether the frequency of each phoneme in different lists is similar across the lists, a one-sample *t*-test was done. In this, the average number of each phoneme across the lists was used as the test value. The results revealed that the frequency of all the phonemes in the language was similar across the lists ($P > 0.05$). The frequency distribution of phonemes in each list was not compared to that in the Tulu language since such data is not available.

Table 2: Mean signal-to-noise ratio-50 and standard deviation of the developed 22 sentence lists

List number	Mean SNR-50±SD
List 1	4.82±1.23
List 2	**4.13±0.93**
List 3	**4.27±1.04**
List 4	**4.33±1.08**
List 5	**−3.83±0.75**
List 6	**−4.03±1.67**
List 7	5.12±1.03
List 8	**−4.32±1.34**
List 9	**−4.4±1.46**
List 10	**−3.85±1.13**
List 11	−3.75±0.90
List 12	−3.77±1.02
List 13	**−4.2±1.55**
List 14	−3.48±1.17
List 15	−4.9±0.94
List 16	**−4.48±0.76**
List 17	**−4.07±0.53**
List 18	−4.87±0.96
List 19	**−4.42±1.11**
List 20	−5.33±0.88
List 21	**−4.15±1.12**
List 22	−3.72±1.21

Lists marked in bold are the final 13 equivalent sentence lists. SD: Standard deviation, SNR: Signal-to-noise ratio

Test reliability

The reliability of SNR-50 obtained from a single participant was determined similarly to that done by Tanniru *et al*.[1] and Nielsen and Dau.[3] For each participant, SNR-50 across 13 lists was averaged and was considered as "true SNR-50." This value of true SNR-50 was deducted from SNR-50 in each list obtained for that participant. These difference values were considered as the deviation of each assessed value of SNR-50 from the true SNR-50. There were 390 values obtained from this analysis (13 lists × 30 participants). Results showed that 71% of the values were within ± 1 dB, 90% of the values were within ±1.5 dB, 94.9% of the values were within ±2 dB, and 99.23% of the values were within ±3 dB of the true SNR-50. The average within-subject SD was 0.99 dB. Thus, the final 13 lists have negligible within-subject deviation, hence, have reasonably good reliability. Furthermore, inter-subject variability was obtained from the same deviations from the true SNR-50 values by averaging the SDs obtained across the participants. This value was 1.04 dB. Thus, within-subject variability was less than between-subject variability.

Significance of signal-to-noise ratio-50 difference across conditions

It is essential to know when the SNR-50 measured across different conditions can be considered significant when they are measured using different equivalent lists. For this, two features need to be considered. As mentioned above, 99.23% of the measured SNR-50 were within ±3 dB of the true SNR-50. The within-subject SD of 0.99 dB was obtained among

Table 3: The average number of words, content words, syllables, phonemes, signal-to-noise ratio-50, and slope of the 13 equivalent lists

Lists	Average number of words per sentence	Average number of content words per sentence	Average number of syllables per sentence	Average number of phonemes per sentence	Average SNR-50 (dB)	Average slope (% per dB SNR)
2	4	3.4	10.1	20.8	−4.13	20.8
3	3.8	3.2	10.2	21.3	−4.27	21.3
4	4	3.3	10.2	22	−4.33	22
5	4	3.3	10.6	22.2	−3.83	22.2
6	3.9	3.2	10	21.3	−4.03	21.3
8	4.3	3.4	10.4	21.7	−4.32	21.7
9	4.2	3.2	10.3	21.1	−4.4	21.1
10	3.7	3.4	9.8	20.2	−3.85	20.2
13	4.2	3.5	10.5	21.3	−4.2	21.3
16	4.2	3.0	10	21	−4.48	21
17	4	3.4	10.5	21.29	−4.07	22.2
19	4	3.5	10.9	21.34	−4.42	22.6
21	4.1	3.6	10.6	21.34	−4.15	21.2

SNR: Signal-to-noise ratio

normal-hearing individuals. If this assumed to be valid for hearing-impaired listeners also, then, the expected deviation of the difference would be ($\sqrt{2} \times 0.99 = 1.40$).[1] When a two-tailed test is considered, the 5% critical region will be limited by 2.74 ($1.96 \times 1.40 = 2.74$). Thus, in a listener, if a difference of 2.75 dB is obtained, for all the practical purposes, it can be taken as an indicator of significant difference among the two conditions in which testing was done.

Discussion

The present study aimed to develop and validate equivalent sentence lists in the Tulu language to assess SRT in noise. The features of developed sentences during phase 1, such as sentences representing daily conversation, free from proper nouns, questions, proverbs, and exclamations, warranted that the sentences received a high rating of naturalness and familiarity, and only a few exclusions were required. The sentences from the first phase were tested on normal hearing individuals at four different SNRs, which helped in estimating the intelligibility of the sentences by using psychometric function. Several previous researchers have used these slopes to make adjustments in SNR such that the intelligibility of the sentences will be equated[3,20] and have found it to be varying from 9% to 17.9% per dB SNR. Using the word scoring method, the mean psychometric slope for 441 sentences ranged between 1.89% and 71.09% per dB SNR in the present study. Adjusting the intelligibility of the sentences using the rescaling method with this wide range of slope would be a tedious task because this requires a large number of participants (e.g.,[12]). Hence, in this study, we used an elimination procedure similar to that used by other investigators,[1,17] which reduces the number of participants studied. In this, the psychometric slopes and SNR-50 were used to select sentences with comparable performance. This allowed us to eliminate sentences that differed significantly from most sentences, that is, more than one SD from the average SNR-50 and psychometric slope.

The average SNR-50 and its SD across the lists obtained in this study were compared to previous studies which have used the sentence scoring method [Table 4]. The SNR-50 in this current research supports the range from previous studies.[16,17,23] The SD of SNR-50 across the lists indicates the variation within the set of lists. A SD of 0.21 dB in the current study suggests a similar level of equivalence across lists. This SD is similar to that obtained by Tanniru *et al*.[1] However, other studies have reported slightly greater SD s ranging from 0.27 to 1.2 dB (mean = 0.77 dB). Thus, the variation across developed lists is low compared to most previous studies, which is a desired feature of the equivalent lists. Removing the nine most divergent lists which had SNR-50 > ±0.5 dB could be one of the reasons to obtain such a low SD. The present study followed the systematic procedure very closely to that followed by Tanniru *et al*.,[1] and it can be seen that both studies obtained similar SDs. The lower variability could even be due to consideration

Table 4: Speech recognition threshold in noise across studies in the literature which have used sentence scoring method compared with the present study

Authors	Overall SRT (dB)	SD
Nilsson *et al*. (1994)[12]	−2.90	0.78
Versfeld *et al*. (2000)[17]	−4.10	0.27
Vaillancourt *et al*. (2005)[20]	−3.30	0.50
Wong and Soli (2005)[16]	−3.90	1.00
Hällgren *et al*. (2006)[22]	−3.0	1.1
Wong *et al*. (2007)[23]	−4.30	0.62
Nielsen and Dau (2009)[3]	−3.15	0.5
Jain *et al*. (2014)[7]	−4.56	0.43
Tanniru *et al*. (2017)[1]	−2.74	0.21
Current study	−4.19	0.21

SD: Standard deviation, SRT: Speech recognition threshold

of added parameters such as the slope of the sentence, number of phonemes, and words in the sentence lists while optimizing the lists.

It is to be noted here that the phoneme frequency distribution data in the Tulu language is not available. Hence, the frequency distribution of phonemes in each list could not be compared to that in the Tulu language. However, this need not be considered a limitation of the developed material, as all the 13 lists have provided equivalent performance. Furthermore, results of a one-sample t-test were for all the phonemes in the language revealed that the frequency of all the phonemes in the language was similar across the lists.

The SD of 0.21 dB also indicates the within-subject variability. Furthermore, within-subject variability was also obtained from the deviations of true SNR-50 (0.99 dB). This within-subject variability was less than that of between-subject variability (1.04 dB), a desirable feature of the standardized speech material. In addition, in a participant, when SNR-50 is measured in two conditions, if they differ by >2.75 dB, that can be considered as a significant difference in SNR-50 between those two conditions. This value of clinical significance is also similar to that obtained by Tanniru *et al*.,[1] who obtained the clinical significance value of SNR-50 as 2.7 dB. Thus, the developed 13 Tulu sentence lists were found to be an equivalent and reliable measure of SNR-50 in native speakers of Tulu.

The findings of the present study are limited to sentence-based scoring. The effect of different types of scoring (word-based and sentence-based) on SNR-50 was not done. Future studies can study the phoneme distribution in the Tulu language and compare the phoneme distribution of present developed lists to that data. Furthermore, the verification of equivalence of the sentence lists was done on 30 participants, including individuals belonging to all the dialects of Tulu. Future studies can investigate the equivalence of the sentence lists across listeners from different dialects separately. In addition, it would be beneficial to assess the equivalency of the developed sentence lists in the hearing-impaired population.

Conclusions

Thirteen sentence lists were developed in the Tulu language, equivalent and comparable to those developed for other languages. The average SNR-50 is −4.19 dB (SD ± 0.21 dB). In addition, within-subject variability on these developed 13 lists was less than that of between-subject variability. Hence, different lists can be used on the same subject during clinical practice and also during research. The difference of mean SNR-50 among the two SNR conditions was 2.75 dB which should be considered as a significant difference while using the developed lists for clinical/research purposes. Thus, the sentence material developed in the Tulu language is highly beneficial in carrying out routine speech audiometry testing of native Tulu speakers with normal hearing sensitivity.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Tanniru K, Narne VK, Jain C, Konadath S, Singh NK, Sreenivas KJ, *et al*. Development of equally intelligible Telugu sentence-lists to test speech recognition in noise. Int J Audiol 2017;56:664-71.
2. Kollmeier B, Wesselkamp M. Development and evaluation of a German sentence test. J Acoust Soc Am 1997;102:2412-21.
3. Nielsen JB, Dau T. Development of a Danish speech intelligibility test. Int J Audiol 2009;48:729-41.
4. Carhart R, Tillman TW. Interaction of competing speech signals with hearing losses. Arch Otolaryngol 1970;91:273-9.
5. Taylor B. Speech-in-noise tests: How and why to include them in your basic test battery. Hear J 2003;56:40-6.
6. Geetha C, Kumar KS, Manjula P, Pavan M. Development and standardisation of the sentence identification test in the Kannada language. J Hear Sci 2014;4:18-26.
7. Jain C, Narne VK, Singh NK, Kumar P, Mekhala M. The Developement of Hindi Sentence Test for Speech Recognition in Noise. Int J Speech Lang Pathol Audiol 2014;2:86-94.
8. Dietz A, Buschermöhle M, Aarnisalo AA, Vanhanen A, Hyyrynen T, Aaltonen O, *et al*. The development and evaluation of the Finnish Matrix Sentence Test for speech intelligibility assessment. Acta Otolaryngol 2014;134:728-37.
9. Hochmuth S, Brand T, Zokoll MA, Castro FZ, Wardenga N, Kollmeier B. A Spanish matrix sentence test for assessing speech reception thresholds in noise. Int J Audiol 2012;51:536-44.
10. Ozimek E, Warzybok A, Kutzner D. Polish sentence matrix test for speech intelligibility measurement in noise. Int J Audiol 2010;49:444-54.
11. Plomp R, Mimpen AM. Improving the reliability of testing the speech reception threshold for sentences. Audiology 1979;18:43-52.
12. Nilsson M, Soli SD, Sullivan JA. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am 1994;95:1085-99.
13. Cameron S, Dillon H. The listening in spatialized noise-sentences test (LISN-S): Test-retest reliability study. Int J Audiol 2007;46:145-53.
14. Avinash, Methi R, Kumar UA. Development of sentence material for Quick speech in noise test (Quick SIN) in Kannada. J Indian Speech Hear Assoc 2009;23:59-65.
15. Killion M, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am 2004;116:2395-405.
16. Wong LL, Soli SD. Development of the cantonese hearing in noise test (CHINT). Ear Hear 2005;26:276-89.
17. Versfeld NJ, Daalder L, Festen JM, Houtgast T. Method for the selection of sentence materials for efficient measurement of the speech reception threshold. J Acoust Soc Am 2000;107:1671-84.
18. Lehiste I, Peterson GE. Vowel amplitude and phonemic stress in american english. J Acoust Soc Am 1959;31:428-35.
19. Ramkissoon I. Speech recognition thresholds for multilingual populations. Commun Disord Q 2001;22:158-62.
20. Vaillancourt V, Laroche C, Mayer C, Basque C, Nali M, Eriks-Brophy A, *et al*. Adaptation of the HINT (hearing in noise test) for adult Canadian Francophone populations. Int J Audiol 2005;44:358-69.
21. Raake A, Katz BF. SUS-based Method for Speech Reception Threshold Measurement in French. Proceedings of the 5th International Conference on Language Resources and Evaluation; 2006. p. 2028-33.
22. Hällgren M, Larsby B, Arlinger S. A Swedish version of the hearing in noise test (HINT) for measurement of speech recognition. Int J Audiol 2006;45:227-37.
23. Wong LL, Soli SD, Liu S, Han N, Huang MW. Development of the Mandarin hearing in noise test (MHINT). Ear Hear 2007;28 Suppl 2:70-4.
24. Theunissen M, Hanekom JJ, Swanepoel D. The development of an Afrikaans test for sentence recognition thresholds in noise. Int J Audiol 2011;50:77-85.
25. Wong LL, Chen Y, Leung KP. The cantonese hearing in noise test for children. Trends Hear 2019;23:1-9.
26. Census I. Abstract of Speakers Strength of Languages and Mother Tongues; 2011. Available from: http://censusindia.gov.in/2011Census/Language-2011/Statement-1.pdf. [Last accessed on 2021 Aug 13].
27. Bhat SN. Descriptive Analysis of Tulu; 1967.
28. Rai VB. Epics in the Oral Genre System of Tulunadu. Vol. 11. Mangalore University, Mangalore: Oral Tradition; 1996. p. 163-72.
29. Kekunnaya KP. A Comparative Study of Tulu Dialects. Udupi, Karnataka: Rashtrakavi Govinda Pai Research Centre; 1994. p. 200.
30. Samuel. Development of SRT Test in Tulu Language. Mangalore Universiity, India; 1998.
31. Carhart R, Jerger J. Preferred method for clinical determination of pure-tone thresholds. J Speech Hear Disord 1959;24:330-45.
32. Dias MA, Devadas U, Rajashekhar B. Development of speech audiometry material in goan Konkani language. Lang India 2015;15:268-80.
33. Francart T, van Wieringen A, Wouters J. APEX 3: A multi-purpose test platform for auditory psychophysical experiments. J Neurosci Methods 2008;172:283-93.

Appendix

List 1: The 10 sentences of List 1 in Kannada and International Phonetic Alphabet (IPA)

Sentence in Kannada	Sentence in IPA	Sentence in English
ಈ ಕೆದುತ ನೀರು ಶೋಕುಂಡು	iː ked̪ut̪ʌ ni:r ʃɔ:kuṇḍu	Water from this lake is good
ಆರ್ದ ಜಿಂಜ ಪಕ್ಕಿಲು ಉಂಡು	a:rḍa dzindzʌ pʌˌkilu uṇḍu	They have many birds
ಎಂಕು ಶೋಕುದ ಗೊಂಬೆ ಬೋಡು	enk ʃɔ:kud̪ʌ gɔmbe bɔ:ḍu	I need a beautiful doll
ಆರೆಗು ಚಳಿಗಾಲೊಡು ರಜೆ ಉಂಡು	a:reg tʃʌḷigʌ:lɔḍu rʌdze uṇḍu	They have winter holidays
ಆಟ ತೂವರೆಗು ಮಾತೆರೆಗ್ಲ ಇಷ್ಟ	a:ṭʌ t̪u:vʌreg mʌ:t̪eregla iʃṭʌ	Everybody likes to watch the dance-drama
ರೈಲುಡು ತಿರ್ಗ್ಯರೆಗು ಎಂಕು ಇಷ್ಟ	railuḍu t̪irgyʌregu enku iʃtʌ	I love traveling in train
ಗಂಟೆ ಪುಚ್ಚೆ ಇಲಿನು ಗಿಡಪುಂಡು	gʌnṭe puˌtʃe ilin giḍʌpuṇḍu	Big cats will chase the rats away
ಆರ್ನ ಇಲ್ಲು ಮಸ್ತು ಗಲೀಜುಂಡು	a:rnʌ iˌlu mʌst̪u gʌli:dzuṇḍu	Their house is very dirty
ಆರು ಯಂಕು ಬಾರಿ ವಿಶೇಷ	a:r jʌnk bʌ:ri viʃe:ʃʰʌ	That person is very special to me
ಆರ್ ಎನ್ನ ಮೋಕೆದೊ ಬಾಲೆ	a:r eˌnʌ mɔ:ked̪ʌ bʌ:le	She/He is my lovely child

IPA: International Phonetic Alphabet