Research Article

A Language Neutral Nonsense Speech-in-Noise Test for Dravidian Language Speakers: Development and Psychometric Evaluation

Prajwal Kumar Eranna[1*], **Jayashree S Bhat**[1], **Aradith S**[2]

[1] *Nitte (Deemed to be University), Nitte Institute of Speech and Hearing, Mangalore, India*
[2] *Shravan Hearing Aid Centre, Shivamogga, India*

Use your device to scan and read the article online

Citation: Eranna PK, Bhat JS, S A. A Language Neutral Nonsense Speech-in-Noise Test for Dravidian Language Speakers: Development and Psychometric Evaluation. Aud Vestib Res. 2026;35(2):171-83.

https://doi.org/10.18502/avr.v35i2.21203

Highlights

- Developed a nonsense SPIN test for Dravidian speakers, independent of native language
- Derived psychometric curve confirms expected non-linear performance with SNR changes
- Demonstrates a moderate to good test-retest reliability with ICC of 0.560–0.830

Article info:
Received: 22 May 2025
Revised: 18 Jul 2025
Accepted: 31 Aug 2025

* **Corresponding Author:**
Nitte (Deemed to be University), Nitte Institute of Speech and Hearing, Mangalore, India.
prajwal03121996@gmail.com

ABSTRACT

Background and Aim: Assessing Speech Perception in Noise (SPIN) in multilingual contexts like India is challenging due to the lack of linguistically appropriate test materials. Recognizing the limitations of existing SPIN tests in multilingual and clinically diverse settings, this study addressed critical need by developing a language-neutral, nonsense SPIN test material tailored for Dravidian languages.

Methods: Nonsense word lists in the consonant vowel consonant vowel format were generated using a random combination of common phonemes in the Dravidian languages (Kannada, Malayalam, Telugu, Tulu, Tamil). These lists were recorded, and the Signal-to-Noise Ratio required to achieve 50% Speech Recognition (SNR50) was used to select optimized lists based on a criterion of mean ±0.15 SD. The final lists were administered to 50 normal-hearing individuals at 0 dB SNR. Language independence was evaluated by comparing performance across speakers of the five languages. Further performance was also assessed across eight SNR levels to establish a psychometric slope function and goodness of fit was assessed. To evaluate test–retest reliability, 12 participants were retested within a one-week interval.

Results: The study resulted in 4 final optimized lists based on SNR50 selection criteria and further analysis. Lists showed sensitivity to varying SNR levels, as reflected by consistent psychometric function slopes. Comparable performance across language groups confirmed the language-independent nature of the test.

Conclusion: Developed test provides audiologists with a reliable and standardized tool to assess SPIN. By eliminating the influence of familiarity and ensuring language neutrality, the test is well-suited for clinical use across speakers of Dravidian languages.

Keywords: Speech perception in noise; development; psychometric function; language independence; slope function; nonsense word

Introduction

In daily life, understanding speech amidst background noise is essential yet challenging. The auditory system relies on mechanisms like auditory closure, binaural interaction, temporal processing, and cognition to manage such conditions. Around 10–15% of people in developed countries have hearing loss [1], which particularly affects Speech Perception in Noise (SPIN) due to factors like reduced audibility, poor frequency and temporal resolution, loss of binaural cues, and increased cognitive load [2]. Even individuals with minimal hearing loss [3] or older adults with normal thresholds often report difficulties in noisy settings, highlighting the role of central auditory and cognitive factors [2, 4]. The SPIN test provides an audiologist a direct measure to assess a patient's ability to understand speech in challenging environments. It aids in diagnosing Central Auditory Processing Disorders [5] and Auditory Neuropathy Spectrum Disorder [6], and can support patient counseling, help establish realistic expectations for hearing aid performance, and aid in tracking the benefits of amplification. The variability in SPIN test scores, even among individuals with normal hearing underscores the significance of the SPIN test [7].

Various SPIN tests exist including the Connected Sentence Test (CST), Hearing in Noise Test (HINT), Words in Noise Test (WIN), Quick Speech-in-Noise Test (QuickSIN), Bamford-Kowal-Bench Speech-in-Noise Test (BKB-SIN), and the Listening in Spatialized Noise-Sentences Test (LiSN-S), etc. [8]. Audiologists often select these tests based on factors such as availability, administration ease, patient age and clinical purpose. A key concern in SPIN testing is the use of linguistically and culturally appropriate materials, as using non-native language stimuli can introduce confounds and may not accurately reflect true speech perception abilities [9]. India, being one of the most linguistically diverse countries in the world, with 121 languages spoken by more than 10,000 people and over 1,600 dialects reported [10], this presents a substantial challenge. While SPIN tests have been developed/adapted in several regional languages in India, such as Tulu [11], Telugu [12], Malayalam [13], Hindi [14], Marathi [15] and Tamil [16], the practical application of these tests in clinical settings can be challenging due to the unavailability of test materials in many regional languages. This poses a challenge in accurately assessing individuals whose native language differs from the test language. While it is ideal to assess speech perception using materials in a person's native language to ensure ecological validity, doing so also introduces linguistic and semantic influences that may confound the results. Specifically, such assessments may not isolate pure phonetic discrimination abilities, as performance could be influenced by factors like lexical familiarity, syntactic expectations, and semantic predictability [17]. Consequently, this makes it difficult to assess speech perception abilities solely at the phonetic or auditory level (bottom-up process), which is often the true focus in audiological evaluations especially while assessing acoustic effects of hearing aid/benefit assessment.

Among the language families in India, the Dravidian languages hold a prominent place, especially in the southern part of the country, and are linguistically, distinct from the Indo-Aryan family (different phonological and syntactic structures), which is predominant in northern India [18]. The major Dravidian languages Telugu, Tamil, Kannada, Malayalam, and Tulu are spoken by a substantial portion approximately 19% of India's population [10]. This vast diversity can be addressed by nonsense word material with their semantic neutrality and allows to control for potential biases introduced by participants' prior knowledge or experience with real words. Such materials have been developed in multiple languages including English [19], Mandarin [20, 21], Greek [22], and German [23]. More recently, Cameron et al. [24] developed the Language-Independent Speech in Noise and Reverberation test (LISiNaR) using Consonant-Vowel-Consonant-Vowel (CVCV) nonsense words, which was validated on second-language English listeners. There exist Phonological differences among different languages [21], which is why even nonsense words have been developed in different languages. Despite being semantically meaningless, nonsense words must still reflect the phonological rules of specific languages, as phonological constraints vary across languages [21].

Considering, audiological practices in India, specifically the southern part of India, audiologists frequently encounter patients from non-native states who belong to a family of Dravidian languages. These languages share similar phonological and syntactic structures [18], making them suitable for a shared nonsense word framework. Thus, developing a Dravidian-language-based nonsense SPIN test could allow audiologists to assess individuals from the southern part of India, independent of their specific native language. Additionally, it also helps in assessing SPIN by minimizing semantic influences and focusing solely on acoustic-phonetic processing (bottom-up). Such a tool would be especially valuable in multi-lingual clinical settings and would offer a standardized approach to evaluating speech perception across diverse linguistic populations. Given the absence of a standardized, language-neutral SPIN test for Dravidian languages, the present study aims to develop such a tool and demonstrate its language independence in individuals aged 18–24 years. The study further aims to evaluate the psychometric function across various Signal-to-Noise Ratio (SNR) levels and examine the test-retest reliability of the developed material.

Methods

A cross-sectional experimental research design has been employed under 2 phases. The first phase of the study includes the development of a test material, followed by establishing a psychometric slope function, demonstrating language independence and assessment of test-retest reliability for the developed material under the second phase.

Phase 1: Development of test material

The study considered the development of the list of nonsense words and it was carried out under the following stages.

Stage I: Formation of nonsense consonant-vowel-consonant-vowel words

The study considered the formation of nonsense words in the CVCV format. Dravidian languages predominantly exhibit simple syllable structures, often conforming to the CVCV pattern and supporting the natural prosody of Dravidian languages, which are typically syllable-timed. This makes the stimuli sound more fluent and native-like, aiding in more ecologically valid assessments of speech perception.

For this purpose, the phonemic inventories of the Dravidian languages were considered and the common phonemes among the Dravidian languages were taken from the study [25]. The Dravidian languages that were considered in the study [25] are Kannada Tamil, Telugu, Malayalam and Tulu. The most common phonemes which are selected from the literature, comprised of 5 vowels and 10 consonants (Table 1). The consonants selected evenly represented the places and manner of articulation. Among vowels, only the short vowels were selected by excluding the long vowels**.** Further, these common phonemes were subjected to all the probable random combinations in the CVCV format using Microsoft Excel software. A customized Microsoft Visual Basic Application (VBA) code was utilized to generate these random combinations, resulting in 2500 probable combinations. Among 2500 combinations, there were 250 combinations starting from each consonant. Further, all the combinations were given to a minimum of 2 native speakers of different languages

Table 1. Selected common phonemes for the purpose of generating nonsense consonant-vowel-consonant-vowel words

	Consonants				
	Bilabials	**Dental**	**Retroflex**	**Palatal**	**Velar**
Stops	/p/	/t/	/θ/	/tiʃ/	/k/
Nasals	/m/	/n/			
Laterals		/l/			
Resonant		/r/			
Semi vowels	/v/				
			Vowel		
	Front			**Back**	
Close	/i/			/u/	
Mid	/e/			/o/	
Open			/a/		

(Kannada, Malayalam, Tamil, Telugu, Tulu) in order to identify any meaningful words that held semantic significance within their respective native languages. The identified meaningful words by the native speakers were excluded, resulting in the removal of 633 meaningful full words out of 2500 random combinations that were generated. Thus, it yielded 1867 nonsense CVCV words, which were utilized to form 30 lists of nonsense words, each containing 25 nonsense words by a random but controlled selection (via MSExcel and VBA). The random selection of nonsense words ensured that each list contained all consonant phonemes and each vowel was equally represented. All selected consonants were equally distributed across each list, with each list containing 25 items, ensuring that all chosen consonant phonemes were present in every list and they were non-repetitive. It was also made sure that the first and the 2nd consonants of the CVCV nonsense words were not the same.

Stage 2: Recording and editing of the nonsense consonant-vowel-consonant-vowel words

The 30 lists of nonsense words were digitally recorded by a female speaker using Adobe Audition® v2.0.5. Recordings were made at a 44.1 kHz sample rate and 24-bit resolution using a condenser microphone connected to a Behringer C-1 preamplifier, placed 30 cm from the speaker's mouth. The selection of the female speaker was based on evaluations of speech rate, suprasegmental features, intelligibility, pronunciation accuracy, and voice quality. Each word was manually extracted from the full recording and saved as a separate .wav file, and normalized for Root Mean Square (RMS) amplitude. Audio quality was evaluated by five audiologists using a 4-point Likert scale (0=not appropriate to 3=totally appropriate) across parameters such as noise, distortion, articulation naturalness, and intonation. Words receiving a score below 2 were re-recorded and re-evaluated using the same procedure. Only those rated ≥2 were retained for further analysis.

Stage 3: Generation of speech spectrum-shaped noise

A customized MATLAB function was utilized to generate speech Spectrum-Shaped Noise (SSN) that has a similar spectral weighting as the nonsense words. This generated noise exhibited a frequency spectrum resembling the Long-Term Average Spectrum (LTAS) of the nonsense words uploaded (Figure 1).

Stage 4: Formation of equally difficult/intelligible lists

Thirty nonsense word lists were evaluated for intelligibility in noise using Signal-to-Noise Ratio required to achieve 50% speech recognition (SNR50) values obtained from five normal-hearing (<15 dB HL) participants. SNR50 was measured using Smriti Shravan software [26] installed in a personal computer routed through an audiometer and the stimulus was

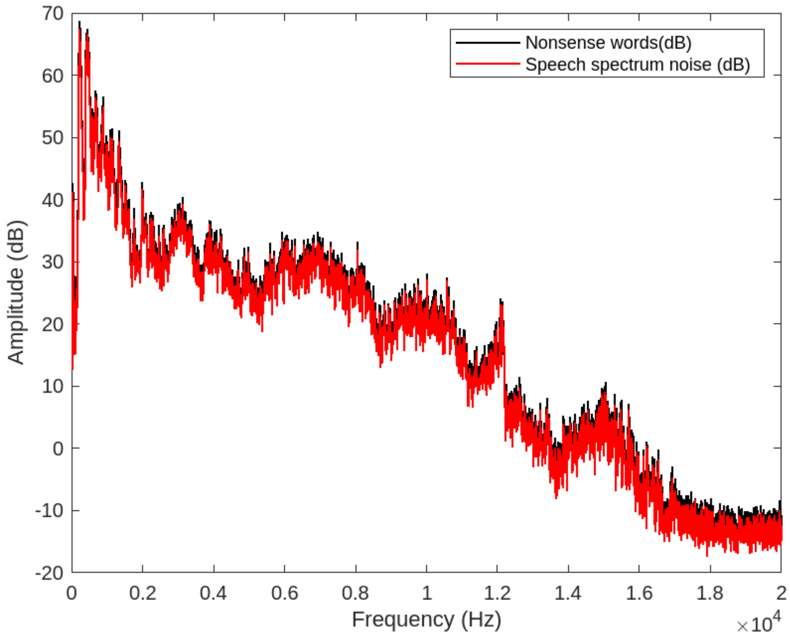

Figure 1. The long-term average spectrum of the speech-shaped noise and concatenated nonsense words

presented using Sennheiser HDA200 headphones. A list of 25 words and SSN generated previously were uploaded and using a 3 down 1 up adaptive procedure, the SNR level was varied in 2 dB step size based on the participant's responses. This procedure involved eight reversals and the midpoint of last 4 reversals were averaged to calculate SNR50. The same procedure was repeated for each list and the presentation order of lists was randomized between the participants. The mean and standard deviation of SNR50 values across the five participants were calculated for all the lists (Figure 2). Only those lists which are having their SNR50 within the 0.15 Standard Deviation (SD) around the mean SNR50 across 5 participants were only selected for further stages as having equal intelligibility.

Stage 5: Mixing nonwords and noise

The final lists of nonsense words were mixed with the speech spectrum noise generated earlier at 0 dB SNR in such a way that there was 500 ms noise present before and after the nonsense word using a customized MATLAB script. This function mixes the speech and noise signals in terms of RMS signal-to-noise ratios.

Phase 2: Establishing a psychometric slope function, demonstrating language independence and assessment of test-retest reliability

Participants

The study recruited 50 adults in the age range of 19–25 years using a purposive sampling. These participants were stratified into groups based on language, with 10 adults representing each language group (Kannada, Malayalam, Tamil, one more language, and Telugu) to demonstrate language independence. All the participants had normal hearing sensitivity ensured by the pure-tone average of <15 dB. All participants' normal outer hair cell functioning was ensured by the presence of transient evoked otoacoustic emissions. Any participants with a risk of auditory processing difficulties indicated by a>6 score in the Screening for Central Auditory Processing in Adults (SCAP-A) checklist [27] were excluded from the study. Any participants with a history of otological, neurological, cognitive complaints, ototoxic drug intake and occupational noise exposure were excluded from the study. All the participants signed written informed consent of willingness.

Procedure

A custom MATLAB script was used to administer the speech-in-noise identification task to establish a psychometric slope function. Those lists that were shortlisted based on 0.15 SD from mean SNR50 and further verified by Friedman's test for their equal performance were only subjected to the assessment of slope function. Participants were presented with a list of 25 recorded nonsense words mixed with SSN at eight different SNRs: –12, –9, –6, –3, 0, +3, +6 and +9 dB. A noise file and 25-word files (.wav format) were selected prior to the experiment. Each word was randomly assigned an SNR, ensuring an even distribution across SNR conditions. For each

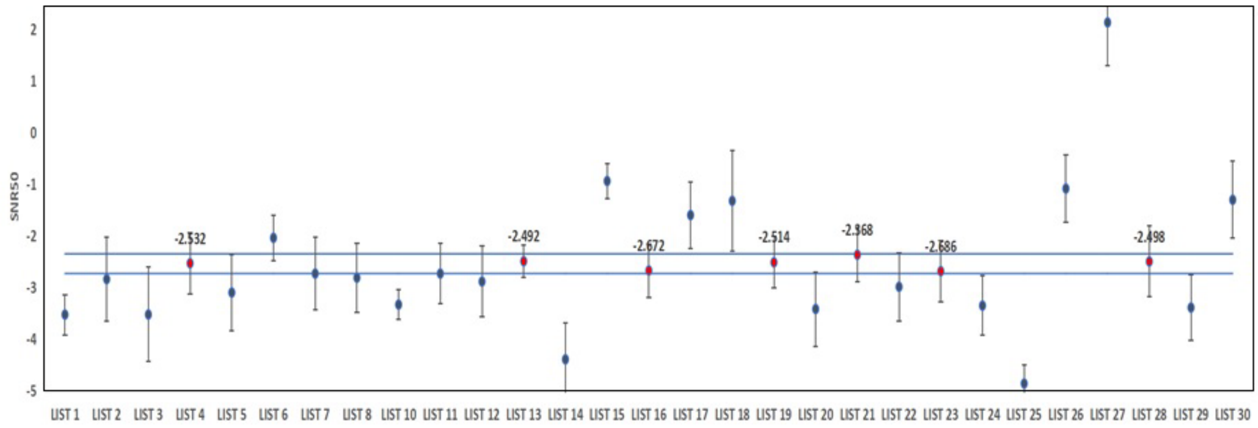

Figure 2. The scatter plot depicting the signal-to-noise ratio required to achieve 50% speech recognition scores for all 30 lists and the two horizontal blue lines indicates the range of –2.35 to –2.73 signal-to-noise ratio required to achieve 50% speech recognition which is the 0.15 SD from a mean signal-to-noise ratio required to achieve 50% speech recognition of –2.54. Those lists whose Signal-to-noise ratio required to achieve 50% speech recognition were within the range are depicted by red dots and were assumed to have equal difficulty and considered for verification of equivalency. SNR50; signal-to-noise ratio required to achieve 50% speech recognition

trial, the selected word was mixed with the noise at the assigned SNR by adjusting the noise amplitude based on the RMS energy of the signals. The mixed stimuli were normalized to prevent clipping and played back to the participant. Thus, each nonsense word in a list was presented at all SNR levels, with the order of presentation randomized across trials. After each presentation, participants were instructed to verbally repeat the word they heard. The clinician conducting the experiment judged the correctness of each response and recorded it using the MATLAB interface. Upon completion of data collection, MATLAB automatically calculated the percentage of correct responses for each SNR level. A logistic regression model was then fitted to the data to derive a psychometric function, and SNR50 was estimated.

Furthermore, to demonstrate language independence, a speech identification task was conducted using premixed word lists at 0 dB SNR through a customized experiment developed in Paradigm Experiment Builder. High-fidelity headphones (Sony MDR-XB450AP) were utilized to deliver the stimuli. Premixed stimuli at 0 dB SNR were presented 65 dB SPL and the participants were instructed to repeat the nonsense CVCV word heard. The examiner will click on the correct or incorrect option based on the response of a participant. Participants' responses were scored dichotomously, with a score of '1' assigned for correct repetitions and '0' for incorrect responses, with a maximum score of 25 for each word list. The order of presentation of wordlists was completely randomized across the participants to avoid the order effect. The same procedure was repeated on 12 of the participants within a one-week interval to assess consistency between scores obtained at different time points.

Results

The data was analyzed using Statistical Package for the Social Sciences (SPSS) software (version 2.0) and Shapiro-Wilks test revealed a non-normal distribution.

Development of nonsense word list

A random combination of the common phonemes resulted in 2500 nonsense CVCV words. 633 out of 2500 were found to be meaningful in either of the 5 Dravidian languages and hence were excluded. Further 30 lists were formed from the rest of the nonwords by a random but controlled selection using Microsoft Excel as explained in the methodology. Descriptive statistics of SNR50 obtained for each of the 30 lists on 5 participants showed that the mean and standard deviation were –2.54 and 1.26 respectively. The SNR50 obtained for all 30 lists is depicted in Figure 2. As discussed in the methodology, mean±0.15 was computed, resulting in a range of –2.35 to –2.73 SNR50. This resulted in 7 lists having their SNR 50 in the above-mentioned range (Figure 2) and they were assumed to have equal difficulty.

Verification of list equivalency/list effect

Descriptive statistics of total correct scores across the 7 lists at 0 dB SNR show that, except for a few, the scores were grossly similar and are depicted in Figure 3 a. The results of Friedman's test for comparison of total scores across 7 lists showed to have a significant list effect ($\chi2(6)=52.162$, $p<0.001$). Thus, Conover's post hoc test was done to check between which list was leading to a significant list effect and it was found that lists 13, 19 and 28 contributed to the significant differences as shown in Table 2. Further, these 3 lists were taken out and Friedman's test was conducted

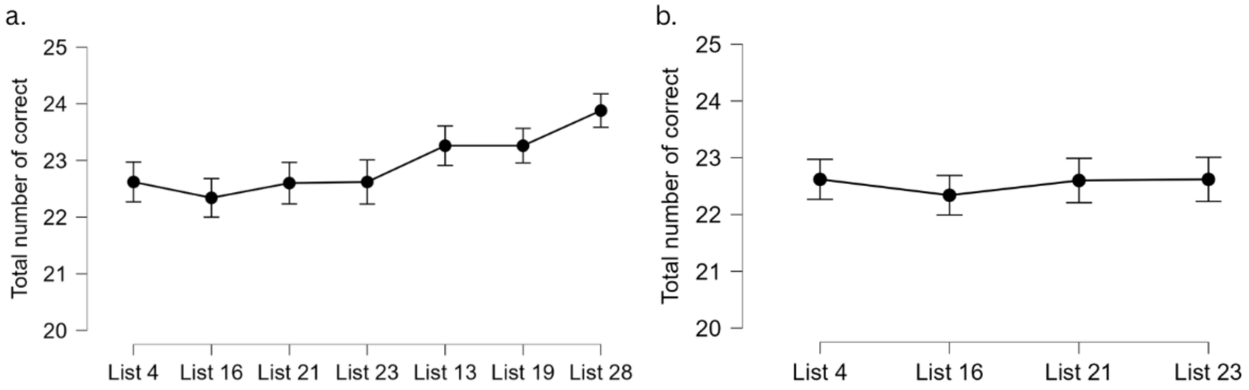

Figure 3. Total correct scores at 0 dB signal-to-noise ratio across (a) 7 lists (b) 4 lists

Table 2. Conver's post hoc test for pairwise comparison of the lists at 0 dB signal-to-noise ratio

Comparison pair		T-stat	p
List 4	**List 16**	0.469	0.640
	List 21	0.888	0.375
	List 23	0.370	0.712
	List 13	2.689	0.008
	List 28	5.354	<0.001
	List 19	2.911	0.004
List 16	**List 21**	1.357	0.176
	List 23	0.839	0.402
	List 13	3.158	0.002
	List 28	5.823	<0.001
	List 19	3.380	<0.001
List 21	**List 23**	0.518	0.605
	List 13	1.801	0.073
	List 28	4.466	<0.001
	List 19	2.023	0.044
List 23	**List 13**	2.319	0.021
	List 28	4.984	<0.001
	List 19	2.541	0.012
List 13	**List 28**	2.665	0.008
	List 19	0.22	0.824
List 28	**List 19**	2.443	0.015

SNR; signal-to-noise ratio

again to compare total scores across the rest of the 4 lists. The results revealed no significant list effect ($\chi2(3)=1.954$, $p=0.582$). The total correct scores of the final 4 lists showing no significant difference are depicted in Figure 3 b.

Psychometric slope curve

Results showed that the mean percent correct scores improved systematically with increasing SNR (–12, –9, –6, –3, 0, +3, +6 and +9 dB) (Table 3) for all 4 lists. A mixed model repeated measure analysis showed a significant main effect of SNR levels ($F_{(7,1372)}=10437.70$, $p<0.001$) whereas no significant main effect ($F_{(3,196)}=1.290$, $p=0.279$) of lists. This indicates that there was no significant difference in the performance across lists irrespective of the SNR levels and there was a significant SNR-dependent improvement in performance irrespective of the list assessed. Further it suggests, that the lists were equivalent in difficulty and responded similarly to changes in SNR. A logistic psychometric function was fit to the group-averaged data using non-linear least mean squares. The model captured the sigmoidal relationship between SNR and speech identification accuracy (Figure 4). The fitted curve also estimated a speech recognition threshold (50% correct point) at –3.14, –2.97, –2.77 and –2.97 dB SNR for lists 21, 23, 16 and 4 respectively. The slope of the curve at these thresholds was consistent with a rapid transition from low to high performance over a narrow SNR range. Model fit was assessed using the coefficient of determination (R^2) and Root Mean Squared Error (RMSE). The results showed that the logistic model provided a good fit to the data, with an R^2 of [0.9986, 0.9983, 0.9966, 0.9973 for List 21, 23, 16 and 4 respectively] and an RMSE of [1.521, 1.711, 2.391, 2.0995 for 21, 23, 16 and 4 respectively], supporting the

Table 3. The mean % correct scores across each signal-to-noise ratio level for the final 4 lists

	List 16		List 21		List 23		List 4	
	Mean (%)	SD	Mean (%)	SD	Mean (%)	SD	Mean (%)	SD
–12 dB SNR	0.24	0.96	0.00	0.00	0.00	0.00	0.00	2.18
–9 dB SNR	5.84	4.21	3.84	3.78	5.04	3.70	6.00	4.70
–6 dB SNR	12.40	5.60	12.50	5.40	12.60	5.50	12.40	5.80
–3 dB SNR	44.08	12.40	49.60	11.71	47.44	12.06	47.92	12.37
0 dB SNR	91.68	5.22	88.72	5.02	89.60	5.36	90.80	5.37
+3 dB SNR	92.88	5.67	90.88	7.79	93.28	5.38	91.52	6.49
+6 dB SNR	96.48	3.58	94.96	2.53	95.60	3.05	95.76	3.73
+9 dB SNR	96.88	2.18	96.32	1.09	96.88	1.67	96.88	2.46

SNR; signal-to-noise ratio

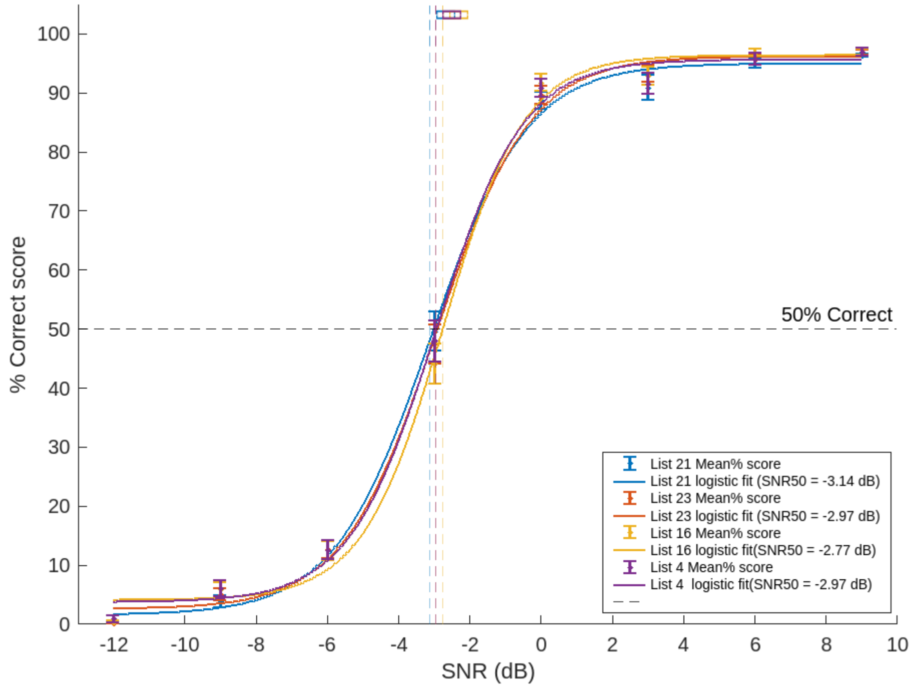

Figure 4. Psychometric slope function of speech identification across different signal-to-noise ratio levels and depicts the signal-to-noise ratio level corresponding to 50% correct identification score for final 4 lists. SNR50; signal-to-noise ratio

appropriateness of the logistic function for describing the data.

Language effect/validation of language independence of the lists

The descriptive statistics showed that the mean scores obtained by the participants of all 5 languages were similar and are depicted in Figure 5. The results of Kruskal Wallis showed no significant differences in the scores obtained by the participants across 5 languages for list 4($\chi^2(4)=4.196$, $p=0.380$), list 16($\chi^2(4)=8.530$, $p=0.074$), list 21($\chi^2(4)=3.282$, $p=0.512$) and list 23($\chi^2(4)=2.693$, $p=0.611$).

Test-retest reliability

The test-retest reliability was assessed using inter-

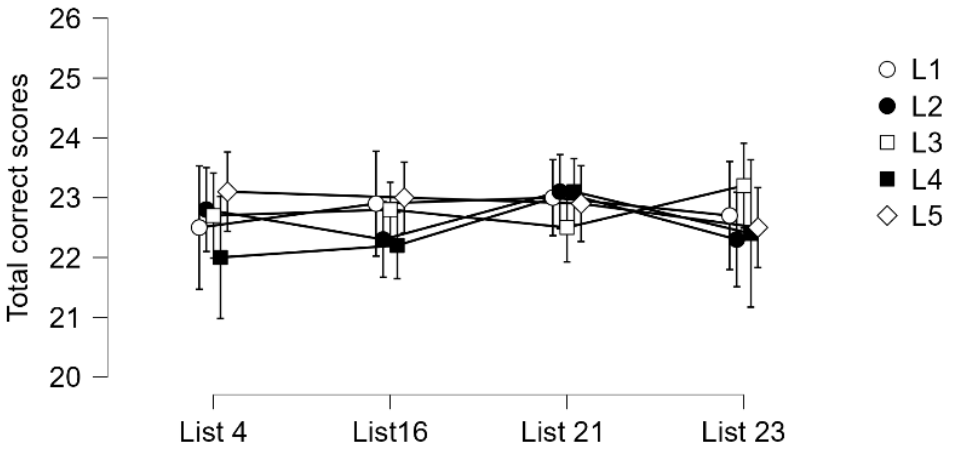

Figure 5. Comparison of the mean correct scores across different language groups for the final 4 lists under 0 dB signal-to-noise ratio. L1, L2, L3, L4 and L5 represent Kannada, Malayalam, Tamil, Telugu and Tulu respectively.

class correlation on 12 participants using scores measured at 2 different time intervals separated by 1 week. The results of descriptive statistics show that the scores are grossly similar under both timelines. The results of the interclass correlation coefficient showed that all 4 lists were found to have a good correlation and it indicated that it ranges from moderate to good reliability across the lists (list 4=0.826, list 16=0.560, list 21=0.796 and list 23=0.830).

Discussion

The primary objective of the present study was to develop and standardize a nonsense SPIN test that can be used across speakers of Dravidian languages, irrespective of their native language. This was achieved by employing nonsense syllables phonotactically valid yet semantically meaningless combinations constructed from phonemes commonly occurring in Dravidian languages. The use of nonsense syllables offers key advantages over real words by minimizing semantic bias, memory effects, and lexical familiarity. Importantly, the constructed syllables were carefully screened to ensure they held no lexical meaning in any of the five major Dravidian languages, thereby supporting their suitability for language-independent assessment. This approach aligns with previous studies that have utilized nonsense syllables to develop linguistically unbiased speech perception tests. For instance, Cameron et al. [24] employed a CVCV syllabic structure, while Ching et al. [28] utilized VCV patterns to achieve language neutrality. Similarly, Kuk et al. [19] adopted CVCV syllables for English, Trimmis et al. [22] implemented CV, VC, and CVC formats in Greek, Schmitt et al. [23] used VCV syllables in German, and Chong et al. [20] applied VCV structures for Mandarin. The choice of CVCV structure in the present study reflects the syllable-timed nature and simple phonotactic patterns typical of Dravidian languages, enhancing ecological validity and ease of processing for native speakers. Thirty lists that were initially formed with a random but controlled selection were subjected to SNR50 calculation to shortlist the number of lists further and ensure equal difficulty across lists. This approach aligns with the methods used in previous studies by Geetha et al. [29], Jain et al. [14], and Bhat et al. [11], who also calculated SNR50 values to select sentences with slopes and SNR50 within ±1 SD on psychometric curves, indicating equal difficulty levels. However, unlike the earlier studies, the present study employed a more stringent cutoff of ±0.15 SD from the mean SNR50. This was deliberately chosen as preliminary analysis showed that relaxing the cutoff would have allowed more than sufficient lists, too at the cost of increased variability. As a result, seven lists in this study achieved SNR50 values within the mean SNR50±0.15 SD range.

Verifying list equivalency was a critical step to ensure that all selected lists maintained comparable difficulty levels. Although initial selection was based on SNR50 values, statistical verification was necessary to confirm functional equivalence. Following the approach of Prasad et al. [13], repeated measures ANOVA with Bonferroni post hoc analysis was conducted, confirming no significant differences among the final four optimized lists. The mean percentage of correct responses for the final four lists ranged from 90.48% to 91.68% at 0 dB SNR. These findings are consistent with Shukla et al. [30], who reported mean scores of 96.16% and 83.33% for English disyllabic words in SSN, and Vineetha et al. [31], who obtained 95.71% accuracy for a Kannada SPIN test using four-talker babble at 0 dB SNR. The slightly lower performance at 0 dB SNR in the current study may be attributed to the use of semantically neutral nonsense words, which lack contextual cues, as well as procedural variations.

The systematic improvement in recognition performance with increasing SNR observed in the current study aligns with well-established auditory perception principles [32]. Results are also in agreement with Shukla et al. [30] who reported a decline in recognition scores with decrease in SNR (–10, –5, 0 and +5 dB SNR). Similar trends were observed by Zhou et al. [33], who evaluated SPIN performance at six SNR levels (20 to –5 dB) in Mandarin. These findings could be attributed to the increased masking. Although the amount of masking depends on the type of noise that is used, as it can lead to different types of masking (energetic and informational). In the current study, SSN was utilized as a masker, which ideally results only in energetic masking. However, the use of other types of maskers, such as speech babble, might lead to a different extent of the SNR effect. [14, 34]. Further, the averaged performance data across seven SNR levels (–12 to +6 dB) demonstrated the expected sigmoidal pattern reflecting progressive unmasking of the speech signal as SNR improved [35]. A logistic psychometric function was fitted using non-linear least squares estimation, accurately modeling the non-linear trajectory of recognition scores. Compared to Prasad et al. [13], the higher mean SNR50 observed in this study likely reflects the absence of semantic cues in the nonsense word material, which increases perceptual difficulty.

Language independence was operationally defined as the absence of significant differences in test scores among Dravidian language groups. Analysis revealed no significant main effect of native language, providing empirical support for the material's language-independent nature. These findings align with Cameron et al. [24], who validated the linguistic neutrality of CVCV nonsense word stimuli by comparing SRTs across speakers of Australian English, Canadian English, and non-native English. Similar to their results, the current study offers converging evidence that phonotactically valid but semantically neutral nonsense words can minimize linguistic bias, making them suitable for cross-linguistic assessment of SPIN.

Conclusion

The present study successfully developed a nonsense speech perception in noise test using phonemes common to Dravidian languages. The finalized material consists of four lists (Appendix A), each containing 25 nonsense words with an equal distribution of shared phonemes. Performance across native speakers of Kannada, Tamil, Telugu, and Tulu showed no significant differences, confirming the test's language-independent applicability within the Dravidian language group. The test demonstrated sensitivity to varying listening conditions through well-defined psychometric functions and exhibited strong test-retest reliability, affirming its consistency and stability. This standardized, semantically neutral tool provides audiologists with a practical method for assessing speech perception without lexical bias.

Ethical Considerations

Compliance with ethical guidelines

The study was conducted in compliance with the ethical guidelines and adhering to the ethical standards of Helsinki. The institutional ethical review board approved the study of K S Hegde Medical Academy, Mangalore, India (EC/NEW/INST/2022/KA/0174). Informed consent was obtained from the participants who agreed to take part in the study.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors' contributions

PKE: Study design, acquisition of data,interpretation of the results, drafting the manuscript and statistical analysis; JSB: Study design, supervision, critical revision of the manuscript; AS: Acquisition of data, drafting the manuscript.

Conflict of interest

No conflicts of interest.

Acknowledgments

The authors would like to thank all the participants who volunteered in the study.

References

1. Kochkin S. MarkeTrak VII: Customer satisfaction with hearing instruments in the digital age. Hear J. 2005;58(9s): 30,32-34,38-40,42-43. [DOI:10.1097/01.HJ.0000286545.33961.e7]

2. Pichora-Fuller MK, Schneider BA, Daneman M. How young and old adults listen to and remember speech in noise. J Acoust Soc Am. 1995;97(1):593-608. [DOI:10.1121/1.412282]

3. Goyal S, Kumar E P, Dutta KK, Varma S G, Thomas P. Audiologists' practice and perspectives on rehabilitation in individuals with minimal-mild degree of hearing loss: a questionnaire-based survey. Egypt J Otolaryngol. 2024;40:28. [DOI:10.1186/s43163-024-00589-2]

4. Humes LE, Dubno JR, Gordon-Salant S, Lister JJ, Cacace AT, Cruickshanks KJ, et al. Central presbycusis: a review and evaluation of the evidence. J Am Acad Audiol. 2012;23(8):635-66. [DOI:10.3766/jaaa.23.8.5]

5. Yathiraj A, Vanaja CS. Criteria to Classify Children as Having Auditory Processing Disorders. Am J Audiol. 2018;27(2):173-83. [DOI:10.1044/2018_AJA-17-0091]

6. De Siati RD, Rosenzweig F, Gersdorff G, Gregoire A, Rombaux P, Deggouj N. Auditory Neuropathy Spectrum Disorders: From Diagnosis to Treatment: Literature Review and Case Reports. J Clin Med. 2020;9(4):1074. [DOI:10.3390/jcm9041074]

7. Wilson RH. Adding speech-in-noise testing to your clinical protocol: Why and how. Hear J. 2004;57(2):10,12,16,19. [DOI:10.1097/01.HJ.0000292386.54654.5d]

8. Wilson RH, McArdle R. Speech signals used to evaluate functional status of the auditory system. J Rehabil Res Dev. 2005;42(4 Suppl 2):79-94. [DOI:10.1682/jrrd.2005.06.0096]

9. Mendel LL. Current considerations in pediatric speech audiometry. Int J Audiol. 2008;47(9):546-53. [DOI:10.1080/14992020802252261]

10. Office of the Registrar General & Census Commissioner, India. Census of India 2011: Abstract of speakers' strength of languages and mother tongues [Internet]. New Delhi: Office of the Registrar General & Census Commissioner; 2018 [cited 2025 Jul 28]. Available from: https://censusindia.gov.in/2011Census/C-16_25062018_NEW.pdf

11. Bhat S, Kalaiah MK, Shastri U. Development and Validation of Tulu Sentence Lists to Test Speech Recognition Threshold in Noise. J Indian Speech Language Hearing Assoc. 2021;35(2):50-6. [DOI:10.4103/jisha.jisha_22_21]

12. Tanniru K, Narne VK, Jain C, Konadath S, Singh NK, Sreenivas KJR, et al. Development of equally intelligible Telugu sentence-lists to test speech recognition in noise. Int J Audiol. 2017;56(9):664-71. [DOI:10.1080/14992027.2017.1307530]

13. Prasad SM, Jain S, Vipin Ghosh PG. Development and Standardization of Sentences for Speech in Noise Test in Malayalam. J India Inst Speech Hearing. 2017;36:48.

14. Jain C, Narne V, Singh NK, Kumar P, Mekhala M. The development of Hindi sentence test for speech recognition in noise. Int J Speech Lang Pathol Audiol. 2014;2(2):86-94.

15. Disha BR, Sharda AS. Development of Speech Perception Test in Noise for Young Marathi-speaking Children. J Indian Speech Language Hearing Assoc. 2024;38(2):122-7. [DOI:10.4103/jisha.jisha_22_21]

16. Gnanasekar S, Vaidyanath R. Perception of Tamil Mono-Syllabic and Bi-Syllabic Words in Multi-Talker Speech Babble by Young Adults with Normal Hearing. J Audiol Otol. 2019;23(4):181-6. [DOI:10.7874/jao.2018.00465]

17. Coene M, Krijger S, Meeuws M, De Ceulaer G, Govaerts PJ. Linguistic Factors Influencing Speech Audiometric Assessment. Biomed Res Int. 2016;2016:7249848. [DOI:10.1155/2016/7249848]

18. Krishnamurti B. The Dravidian Languages. Cambridge: Cambridge University Press; 2003.

19. Kuk F, Lau CC, Korhonen P, Crose B, Peeters H, Keenan D. Development of the ORCA nonsense syllable test. Ear Hear. 2010;31(6):779-95. [DOI:10.1097/AUD.0b013e3181e97bfb]

20. Chong FY, Cheoy LP, Mazlan R, Maamor N. Performance-intensity functions of Mandarin fricative-affricate nonsense word test: preliminary findings. Speech Lang Hear. 2020;23(3):121-32. [DOI:10.1080/2050571X.2019.1576364]

21. Cheoy LP, Chong FY, Mazlan R, Lim HW. Development of the Mandarin Nonsense Word Identification Test. Int J Audiol. 2021;60(8):578-87. [DOI:10.1080/14992027.2020.1864485]

22. Trimmis N, Vrettakos G, Gouma P, Papadas T. Speech Audiometry: Nonsense Monosyllabic Lists in Modern Greek. J Hear Sci. 2012;2(3):41-9. [DOI:10.17430/883512]

23. Schmitt N, Winkler A, Boretzki M, Holube I. A Phoneme Perception Test Method for High-Frequency Hearing Aid Fitting. J Am Acad Audiol. 2016;27(5):367-79. [DOI:10.3766/jaaa.15037]

24. Cameron S, Boyle C, Dillon H. The development of the Language-Independent Speech in Noise and Reverberation test (LISiNaR) and evaluation in listeners with English as a second language. Int J Audiol. 2023;62(8):756-66. [DOI:10.1080/14992027.2022.2078432]

25. Sjoberg AF. The Structure of Dravidian Languages. Final Report.; 1971. https://eric.ed.gov/?id=ED055516

26. Kumar AU, Sandeep M. Auditory Cognitive Training Module. ARF-Funded Departmental Project Submitted to All India Institute of Speech and Hearing; 2013.

27. Vaidyanath R, Yathiraj A. Screening Checklist for Auditory Processing in Adults (Scap-A): Development and Preliminary Findings. J Hear Sci. 2014;4(1):27-37. [DOI:10.17430/890788]

28. Ching TYC, Dillon H, Hou S, Seeto M, Sodan A, Chong-White N. Development and Evaluation of a Language-Independent Test of Auditory Discrimination for Referrals for Cochlear

Implant Candidacy Assessment. Ear Hear. 2022;43(4):1151-63. [DOI:10.1097/AUD.0000000000001166]

29. Geetha C, Kumar KSS, Manjula P, Pavan M. Development and Standardisation of the Sentence Identification Test in the Kannada Language. J Hear Sci. 2014;4(1):18-26. [DOI:10.17430/890267]

30. Shukla B, Rao BS, Saxena U, Verma H. Measurement of Speech in Noise Abilities in Laboratory and Real-World Noise. Indian J Otol. 2018;24(2):109. [DOI:10.4103/indianjotol.INDIANJOTOL_134_17]

31. Vineetha CV, Suma R, Nair SP. Effect of Bilingualism on Speech in Noise Perception in Young Adults. Language in India. 2013;13(6):799.

32. Plomp R. A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. J Speech Hear Res. 1986;29(2):146-54. [DOI:10.1044/jshr.2902.146]

33. Zhou R, Zhang H, Wang S, Chen J, Ren D. Development and Evaluation of the Mandarin Quick Speech-in-Noise Test Materials in Mainland China. J phonet Audio. 2017;2(3). [DOI:10.4172/2471-9455.1000124]

34. Stuart A, Phillips DP. Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbyacusic listeners. Ear Hear. 1996;17(6):478-89. [DOI:10.1097/00003446-199612000-00004]

35. 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(2):1085-99. [DOI:10.1121/1.408469]

Appendix A. Language independent test material

S.N	LIST 1	LIST 2	LIST 3	LIST 4
1.	/riru/	/vane/	/pivu/	/pavi/
2.	/raʈe/	/voɭo/	/peku/	/poma/
3.	/rine/	/veka/	/poʈo/	/pupa/
4.	/vemi/	/ravu/	/ʈunu/	/ʈat̪o/
5.	/voɭu/	/ruɭo/	/ʈeʈo/	/ʈatʃu/
6.	/vima/	/rotʃa/	/ʈake/	/ʈuko/
7.	/ɭiʈo/	/ɭavo/	/tʃopo/	/tʃina/
8.	/ɭape/	/ɭuvi/	/tʃat̪i/	/tʃari/
9.	/ɭiva/	/ɭake/	/tʃeto/	/tʃuka/
10.	/nope/	/nome/	/t̪it̪e/	/t̪it̪a/
11.	/nuʈi/	/nuka/	/t̪ema/	/t̪amo/
12.	/natʃe/	/natʃi/	/t̪onu/	/t̪eke/
13.	/meɭo/	/met̪e/	/kama/	/kopo/
14.	/mitʃe/	/mavi/	/kotʃu/	/ket̪i/
15.	/mat̪u/	/mira/	/kuʈa/	/kavo/
16.	/kuki/	/kave/	/miɭe/	/maru/
17.	/kovu/	/kuro/	/movo/	/mevi/
18.	/tʃapa/	/tʃitʃu/	/niʈo/	/napo/
19.	/tʃot̪i/	/tʃeɭo/	/nami/	/neʈu/
20.	/t̪eɭo/	/t̪oni/	/ɭeɭo/	/ɭut̪e/
21.	/t̪amu/	/t̪ipe/	/ɭatʃa/	/ɭotʃa/
22.	/ʈavo/	/ʈomu/	/rime/	/rako/
23.	/ʈire/	/ʈuva/	/retʃo/	/reɭi/
24.	/peni/	/poʈe/	/vape/	/vopa/
25.	/pame/	/puku/	/vumi/	/vut̪i/