Auditory and Vestibular Research

Comparison of whole-word and phoneme scoring for SBMU-1 monosyllabic words in presence of speech-spectrum noise: A psychometric function study

Zahra Moeinfard
Department of Audiology, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Tehran, Iran
Orcid: 0000-0003-3737-3794
e-mail: moeinfard.zahra@gmail.com

Mohammad Ebrahim Mahdavi
Department of Audiology, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Tehran, Iran
Orcid: 0000-0002-7589-7065
e-mail: mahdavime@sbmu.ac.ir

Hamid Jalilvand
Department of Audiology, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Tehran, Iran
Orcid: 0000-0002-2351-5918
e-mail: hamidjalilvand@sbmu.ac.ir

Highlights
Phoneme scoring gave higher recognition scores, than whole-word scoring in most SNRs
SBMU-1 lists showed similar psychometric functions, confirming equivalency in noise
Whole-word scoring is more natural, while phoneme scoring is more sensitive in noise

Abstract
Background and Aim: Speech audiometry assesses functional hearing beyond pure-tone thresholds, reflecting real-world speech perception. The present study compared whole-word and phoneme scoring methods for the Persian SBMU-1 monosyllabic word lists presented in speech-spectrum noise to determine their psychometric equivalence and sensitivity to signal-to-noise ratio (SNR) changes.
Methods: Twenty-two young adults with normal hearing participated. SBMU-1 consonant–vowel–consonant (CVC) words were presented binaurally at six SNRs (−5, 0, +5, +10, +15, +20 dB) in speech-spectrum noise at 60 dB HL. Recognition performance was analyzed using whole-word and phoneme scoring. List equivalency and scoring effects were examined using Friedman and Wilcoxon signed-rank tests with Bonferroni correction.
Results: Speech recognition improved systematically with increasing SNR for both scoring methods. Phoneme scoring consistently yielded higher scores than whole-word scoring, especially under low SNRs, revealing a 10–20% performance advantage. At high SNRs, the two methods converged. Across lists, differences were minor and list-specific, confirming the general equivalency of SBMU -1 lists in noise.
Conclusion: Phoneme scoring provides a more sensitive measure of perceptual performance in noise by capturing partial recognition, whereas whole-word scoring better represents functional communication ability. The SBMU-1 word lists are psychometrically balanced and suitable for both clinical and research applications in Persian speech audiometry.
Keywords: Speech audiometry, word recognition, phoneme scoring, Persian SBMU -1 words, speech perception in noise

Introduction
Speech audiometry is a cornerstone of audiological assessment, providing critical information about an individual's ability to perceive and understand speech. Assessment of hearing ability extends beyond the measurement of pure-tone thresholds, as the ability to detect sounds does not necessarily reflect how effectively

an individual can understand speech in daily communication [1]. When speech stimuli are employed to evaluate central auditory processing, they engage broader domains in the auditory system. The linguistic, phonetic, and syntactic features inherent in speech make the auditory processing of such signals more complex than that involved in pure-tone assessment [2]. In audiology, one of the most widely used clinical tools to evaluate speech perception is the Word Recognition Score (WRS), sometimes referred to as speech discrimination score. The importance of WRS lies in its clinical utility: it helps differentiate between conductive and sensorineural hearing loss, evaluates the impact of retrocochlear pathology, and guides audiologists in counseling, rehabilitation planning, and candidacy determination for hearing aids or cochlear implants. Unlike pure-tone audiometry, which only quantifies audibility, WRS reflects functional communication ability, making it crucial for predicting real-world listening performance. Furthermore, reduced word recognition despite adequate audibility often indicates auditory distortion or neural involvement, underscoring its role in diagnostic audiology [1,2].
For decades, the standard method for quantifying this ability has been whole-word scoring, wherein a listener's response is marked as correct only if the entire word is accurately repeated. While this method is intuitive and widely used, it has inherent limitations, including a relatively small number of data points per test list and potential influences from cognitive, linguistic, and contextual factors—so-called "top-down" processing. In response to these limitations, phoneme scoring has been proposed as a viable alternative [3-6].
The fundamental relationship between these two scoring methods was established in the seminal work of Markides, who systematically compared whole-word and phoneme scoring using five and ten phonemically balanced (PB) word lists, presented via headphones and in the free field, respectively. His findings demonstrated a consistent and substantial difference: phoneme scores were approximately 20–25% higher than whole-word scores across the linear portion of the speech discrimination curve. When translated into intensity levels, this difference amounted to a 4–5 dB increase in sensitivity for the phoneme scoring method, a result observed consistently across listeners with normal hearing, conductive hearing loss, and sensorineural hearing loss [6].Building upon this foundation, Billings et al. sought to refine phoneme scoring for modern clinical and research applications, particularly in the context of speech-in-noise testing—a scenario that better reflects the real-world listening challenges reported by many individuals, especially those with hearing impairment [3]. Their study had three primary aims: to establish a standardized set of phoneme scoring guidelines to minimize tester bias, to apply these guidelines in testing older adults and individuals with hearing loss, and to determine whether the scoring method (word vs. phoneme) modifies the measured effects of age and hearing impairment. The results of Billings et al. confirmed that phoneme scores were notably higher (by about 10–30%) than word scores [3].
The SBMU monosyllabic words (SBMU -1 Words) have recently been developed with the aim of creating a psychometrically homogeneous word lists for the Persian [7]. These consonant–vowel–consonant (CVC) () words are organized into three 50-word lists, with each full list comprising two 25-word sub-lists. The present study was conducted to compare whole-word scoring with phoneme scoring during the presentation of these words in the presence of standard speech-spectrum noise.

Methods

Participants
The participants were 22 undergraduate and graduate students from the School of Rehabilitation Sciences at SBMU. Their average age was 23 years. The participants' hearing levels were below 15 dB in the 250-8000 Hz frequency range. Inclusion criteria included no history of ear or neurological disorders and no use of alcohol or drugs.

Measures
SBMU -1 words, all following the CVC phonetic cluster, were used as target words. In the development of the word lists of SBMU -1, criteria such as word frequency, common usage and familiarity, and psychometric homogeneity were applied. The selection of words for the SBMU-1 lists was performed according to Wilson and Carter to ensure that the average difficulty level of the lists was similar [10]. Each word was mixed with a standard speech-spectrum noise segment with a 600-millisecond duration at five different signal-to-noise ratios (-5, 0, 5, 10, 15, and 20 dB). Five randomized sets of words were constructed and presented binaurally via

headphones to the participants at an intensity level of 60 dB HL. An Interacoustics AD229 audiometer was used, and participants attended sessions lasting 45–50 minutes, during which they received 900 presentations (6 signal-to-noise ratios multiplied by 150 words). For half of the participants, signal-to-noise ratios (SNRs) were presented in ascending order, and for the other half, in descending order.

Statistical analysis

Considering a Type I error of $\alpha = 0.05$, a Type II error of $\beta = 0.2$ (i.e., a power of 80%), and conducting a pilot study on 5 participants, the sample size was calculated using the above formula, resulting in a total of 25 participants. Due to the non-normal distribution of the residuals, nonparametric statistical tests were used to compare the scoring methods and the lists in terms of mean recognition scores. The psychometric function was fitted using a third-degree polynomial function.

Results

WRS increased progressively with higher signal-to-noise ratios for both whole-word and phoneme scoring across all three test lists. At the lowest SNR (-5 dB), participants showed poor performance, with mean whole-word scores ranging from 28.61% to 33.42% and phoneme scores ranging from 43.91% to 45.18%, reflecting significant difficulty in understanding speech in challenging listening conditions. As SNR improved, both scoring methods demonstrated substantial gains; at 0 dB, mean scores rose to 65–73% for whole-word scoring and 78–83% for phoneme scoring, while at 20 dB, performance approached ceiling levels, with whole-word means exceeding 97% and phoneme means exceeding 98%. Across all SNRs, phoneme scoring consistently yielded higher scores than whole-word scoring, indicating that listeners were often able to correctly perceive individual phonemes even when entire words were not recognized. Variability, as indicated by standard deviation, was highest at low SNRs and decreased at higher SNRs, suggesting more consistent speech perception under favorable listening conditions. These patterns were consistent across all three test lists, demonstrating the reliability of the measures and the sensitivity of WRS measured using SBMU -1 words to varying noise levels.

To examine the differences in performance across the three lists under various SNR conditions, the Friedman test was applied (a nonparametric test for repeated measures). Where significant differences were found, pairwise comparisons with Bonferroni correction were planned. The significance level was set at $\alpha = 0.05$, which after Bonferroni adjustment was reduced to 0.017.

Whole-word scoring

The Friedman test was applied to compare recognition performance across the three monosyllabic word lists under six SNR conditions (Table 1). At SNR = –5 dB, the result approached significance ($\chi^2 = 6.52$, $p = 0.038$), but it did not survive Bonferroni correction (adjusted $\alpha = 0.017$). Similarly, at +15 dB, no significant difference was observed among lists ($\chi^2 = 5.95$, $p = 0.051$).

In contrast, significant differences were found at SNRs of 0, +5, +10, and +20 dB. Specifically, at 0 dB ($\chi^2 = 10.50$, $p = 0.005$), post-hoc comparisons indicated that List 2 yielded higher scores than List 1, while differences between Lists 2 and 3, and Lists 1 and 3, were smaller and less consistent. At +5 dB ($\chi^2 = 15.24$, $p < 0.001$), List 2 produced significantly lower scores than both Lists 1 and 3, suggesting greater relative difficulty. At +10 dB ($\chi^2 = 11.18$, $p = 0.004$), List 1 produced lower scores compared to Lists 2 and 3, indicating that it was slightly more challenging at this SNR. Finally, at +20 dB ($\chi^2 = 8.49$, $p = 0.014$), the main difference was between List 2 and List 3, with List 2 producing marginally higher scores.

Phoneme scoring

The analysis was conducted at six SNR levels ranging from –5 to +20 dB (Table 2). At SNR = –5, +5, +10, +15, and +20, the p values (0.554, 0.217, 0.154, 0.174, and 0.079, respectively) were all greater than 0.05. Thus, no significant differences were observed among the three lists in these conditions. At SNR = 0, the Friedman test yielded $p = 0.024$. While this result was significant at the 0.05 level, it did not reach the stricter threshold of $p < 0.017$ after Bonferroni correction. Therefore, the difference was not considered statistically significant. The findings indicate that across most SNR conditions (both negative and positive), the three test lists produced comparable performance, with no statistically significant differences. A trend toward divergence was observed at SNR = 0, but this did not survive correction for multiple comparisons, suggesting it may have been a chance finding. Overall, the three lists can be considered equivalent in noise considering Phoneme scoring. The results demonstrate that, while the three lists are largely comparable under very difficult (–5 dB) and relatively easy (+15 dB) listening conditions, systematic differences emerge at intermediate SNRs. These differences appear to be list-specific rather than global. For example, List 1 was more difficult at 0 and +10 dB, whereas List 2 was more challenging at +5 dB. At +20 dB, List 3 produced slightly lower recognition scores compared to List 2, though all scores were close to ceiling levels.

Whole-word versus phoneme scoring

The Wilcoxon signed-rank test revealed significant differences between phoneme scoring and whole-word scoring across nearly all lists and SNR conditions. For Lists 1 and 2, phoneme scoring consistently produced higher recognition scores than whole-word scoring at all SNRs (all $p < 0.05$). For List 3, significant differences were observed from –5 to +15 dB ($p < 0.01$), but at +20 dB the difference was not significant ($p = 0.292$). The nalysis indicates a systematic advantage of phoneme scoring over whole-word scoring across nearly all SNRs and lists. This effect is strongest in challenging listening conditions (–5, 0, +5 dB), where phoneme-level partial recognition still contributes to performance, but whole-word scoring penalizes partial errors.

Psychometric function characteristics

The characteristics of the psychometric functions and the predicted SNR and slope values for the two scoring methods are presented in Table 3 and Figure 1.

The predicted signal-to-noise ratio (SNR) for different recognition levels of SBMU -1 words show distinct patterns for phoneme-based and whole-word scoring. For phoneme scoring, the SNR required to reach 50% recognition is approximately –4.5 dB across all three lists, with the slope between 20% and 80% recognition ranging from 7.2 to 8.4%/dB, indicating a relatively steep improvement in recognition as SNR increases. In contrast, for whole-word scoring, the SNR at 50% recognition is slightly higher, ranging from –2.4 to –3.1 dB, while the slope is moderately lower (6.3–7.5%/dB), suggesting a slower rise in recognition with increasing SNR. Notably, List 2 exhibits the steepest slopes in both scoring methods, indicating higher sensitivity to SNR changes. These results highlight that phoneme-level recognition is generally achieved at lower SNRs with sharper transitions, whereas whole-word recognition requires higher SNRs and shows more gradual improvement, reflecting differences in perceptual difficulty and scoring granularity between the two methods.

Distribution of consonants and vowels across the lists of SBMU -1 words The frequency counts of consonants, categorized by manner of articulation, and the vowels present in the three **SBMU-1 monosyllabic word lists** (List 1, List 2, List 3) are presented in Table 4. Overall, the distribution of consonant manners of articulation was relatively consistent across the three lists. Stops and Fricatives were the most abundant categories in each list, with counts remaining high and stable (Stops: List 1=33, List 2=31, List 3=30; Fricatives: List 1=32, List 2=33, List 3=28). Nasals showed a noticeable increase from List 1 (11) to List 3 (18), while Approximants decreased slightly from List 1 (20) to List 3 (16). Affricates were the least common consonant category but were present in all lists, with the highest count in List 2 (7). The vowel distribution exhibited more notable variation across the lists. The vowels /æ/ and /ɑ/ were among the most frequent, though /æ/ decreased from 14 instances in List 1 to only 7 in List 3. Conversely, the use of diphthongs and back vowels showed some shifts; the diphthong /oʊ̯/ was absent in List 1 but appeared 2 and 4 times in List 2 and List 3, respectively. The back vowel /o/ was least frequent in List 2 (3) but more common in List 1 (5) and List 3 (7). The vowel /u/ was most frequent in List 1 and List 2 (12 each) and less so in List 3 (8).

Discussion

The present study sought to compare whole-word and phoneme scoring methods for the Persian SBMU -1 monosyllabic word lists presented in speech-spectrum noise. Our findings align with and extend the existing body of literature, demonstrating a systematic and significant advantage for phoneme scoring across nearly all tested SNRs.

The most consistent finding was that phoneme scores were markedly higher than whole-word scores, particularly under adverse listening conditions (SNRs from -5 to +5 dB). At the most challenging SNR of -5 dB, the mean phoneme score was approximately 15% higher than the whole-word score. This disparity underscores a fundamental difference between the two scoring methods: phoneme scoring credits partial word recognition, allowing listeners to score points for correctly identified phonemes even when the entire word is misperceived. In contrast, whole-word scoring imposes an all-or-nothing criterion, which more heavily penalizes listeners in conditions where the acoustic signal is degraded. This result is in direct agreement with the foundational work of Markides [6], who reported a 20-25% performance advantage for phoneme scoring in the linear portion of the discrimination curve, a finding that has been consistently replicated in subsequent studies [3,4].

The convergence of the two scoring methods at the most favorable SNR (+20 dB) is also consistent with theoretical expectations. As the speech signal becomes fully audible above the noise floor, listeners are increasingly able to correctly identify entire words, minimizing the opportunity for partial credit to influence the final score. This convergence at ceiling performance suggests that while phoneme scoring is a more sensitive tool for quantifying performance in noise, whole-word scoring remains a valid indicator of optimal speech recognition ability in quiet or near-quiet conditions.

Beyond the simple comparison of scores, our study offers insights into the clinical utility of each method. The superior sensitivity of phoneme scoring in noise makes it a valuable tool for detecting subtle deficits in auditory processing that may be masked by whole-word scoring in conventional tests. This is particularly relevant for populations with sensorineural hearing loss, who often report significant difficulty understanding speech in noisy environments despite relatively preserved quiet word recognition. As Billings et al. argued, phoneme scoring can provide a more precise measure of the peripheral and central auditory system's ability to resolve acoustic-phonetic details under challenging conditions [3]. Furthermore, by tripling the number of scorable items per word list (from 50 words to 150 phonemes), phoneme scoring enhances the statistical reliability of the test, a point robustly demonstrated by Gelfand with his Computer Assisted Speech Recognition Assessment (CASRA) paradigm [4].

The results demonstrate that, while the three word lists are largely comparable under very difficult (–5 dB) and relatively easy (+15 dB) listening conditions, systematic differences emerge at intermediate SNRs. These differences appear to be list-specific rather than global. For example, List 1 was more difficult at 0 and +10 dB, whereas List 2 was more challenging at +5 dB. At +20 dB, List 3 produced slightly lower recognition scores compared to List 2, though all scores were close to ceiling levels. Such variability across lists is not unexpected in speech audiometry materials. Subtle differences in lexical familiarity, phonemic balance, or acoustic properties can lead to small but measurable discrepancies in recognition performance when listening conditions are neither floor nor ceiling. Importantly, the observed differences were relatively modest in size (generally 3–8%) and did not consistently favor a single list across all conditions. Methods to standardize speech audiometry materials and reduce variability often involve careful development and psychometric evaluation of the test materials [8,9] include homogeneity of stimulus items [10], combining lists [11], statistical validation [9], digital intensity adjustment [8].

Researchers hold divergent views regarding the effect of noise on phoneme perception. Some studies suggest that nasals and approximants are more susceptible to degradation than fricatives and stops [12]. Parikh and Loizou reported that stop consonants were still recognized accurately even at −5 dB SNR, implying that listeners relied on cues beyond the burst—such as formant transitions—when noise masked the primary acoustic signal [13]. In contrast, Nishi et al. found that stop consonants were the most difficult phoneme type for both children and adults to perceive in noise [14]. As shown in Table 4, the three lists maintain a similar and robust distribution of consonants, especially stops and fricatives, ensuring these common consonant types are well-represented.

Overall, these findings suggest that the three lists are reasonably balanced and suitable for use in research and clinical contexts. However, minor differences in list difficulty at specific SNRs should be acknowledged, particularly when precise comparisons across conditions are required.

Limitations and Future Directions

This study was conducted with young adults with normal hearing. Future research should extend these findings to broader clinical populations across different age groups, including individuals with sensorineural hearing loss of varying degrees and configurations. It would be particularly insightful to investigate the relationship between phoneme error patterns—regarding place and manner of articulation—and specific audiometric profiles, which could inform more targeted auditory rehabilitation strategies. The SBMU -1 word lists were developed with psychometric homogeneity in order to guarantee similar average difficulty across lists; however, no digital intensity adjustment of the Root mean square (RMS) values was carried out, which could be considered in the future, more precise studies. The study did not account for the right-ear advantage or the possible effect of ear dominance, which could have considered as intervening variables. Furthermore, comparing the ×××-1 lists with other standardized Persian materials, including nonsense sentences and phonemic structures such as consonant-vowel (CV) and consonant–vowel–consonant-consonant (CVCC), using both scoring methods would provide additional evidence for their clinical validity. Minor variations in list difficulty observed at specific SNRs should be considered when making precise condition comparisons. Future studies may consider refining these lists based on detailed participant score analyses to develop more equivalent versions that are more suitable for clinical use.

Conclusion

In summary, this study confirms that scoring method significantly influences measured speech recognition performance in noise. Phoneme scoring for the SBMU -1 word lists provide a more sensitive and reliable measure of auditory perception in noisy environments by capturing partial word recognition, whereas whole-word scoring reflects a stricter, functionally oriented criterion. The demonstrated equivalency of the SBMU -1 lists validates their use in both clinical and research settings. Audiologists and researchers are encouraged to consider their specific objectives when choosing a scoring method, and may benefit from reporting both scores to gain a comprehensive understanding of a listener's speech perception abilities.

Ethical considerations

Compliance with ethical guidelines

This study was approved by the Research Ethics Committee of Shahid Beheshti University of Medical Sciences (IR.SBMU.RETECH.REC. 1404.028).

Funding

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

Authors' contributions

ZM: Study design and acquisition of data; drafting the manuscript; MEM: Resaerch idea and supervision, Statistical analysis, interpretation of the results, and critical revision of the manuscript; HJ: Interpretation of the results, and validation the final revision of the manuscript. All authors discussed the results and contributed to the final manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

The authors express their sincere appreciation to all the students of the School of Rehabilitation Sciences at Shahid Beheshti University of Medical Sciences for their voluntary participation in this study. Their valuable cooperation and enthusiasm greatly contributed to the successful completion of this research.

References

1. Lawson GD, Peterson ME. Speech Audiometry. San Diego: Plural Publishing, Inc.; 2011.
2. Ashrafi M, Rezaei Sakha M. Development of a Binaural Masking Level Differences Test Application Using Tonal and Speech Stimuli in Persian and Assessing Its Test-Retest Reliability in Normal-Hearing Young People. Aud Vestib Res. 2022;31(4):311-8. [DOI:10.18502/avr.v31i4.10736]
3. Billings CJ, Penman TM, Ellis EM, Baltzell LS, McMillan GP. Phoneme and Word Scoring in Speech-in-Noise Audiometry. Am J Audiol. 2016;25(1):75-83. [DOI:10.1044/2016_AJA-15-0068]
4. Gelfand SA. Optimizing the reliability of speech recognition scores. J Speech Lang Hear Res. 1998;41(5):1088-102. [DOI: 10.1044/jslhr.4105.1088]
5. Gelfand SA. Tri-word presentations with phonemic scoring for practical high-reliability speech recognition assessment. J Speech Lang Hear Res. 2003;46(2):405-12. [DOI: 10.1044/1092-4388(2003/033)]
6. Markides A. Whole-word scoring versus phoneme scoring in speech audiometry. Br J Audiol. 1978;12(2):40-6. [DOI: 10.3109/03005367809078852]

7. Mahdavi ME, Rabiei A. Psychometric function characteristics of Persian consonant-vowel-consonant words. Aud Vestib Res. 2021;30(1):50-55. [DOI:10.18502/avr.v30i1.5311]
8. Nissen SL, Harris RW, Channell RW, Conklin B, Kim M, Wong L. The development of psychometrically equivalent Cantonese speech audiometry materials. Int J Audiol. 2011;50(3):191-201. [DOI: 10.3109/14992027.2010.542491]
9. Xi X, Wang Y, Shi Y, Gao R, Li S, Qiu X, et al. Development and Validation of a Mandarin Chinese Adaptation of AzBio Sentence Test (CMnBio). Trends Hear. 2022;26:23312165221134007. [DOI: 10.1177/23312165221134007]
10. Wilson RH, Carter AS. Relation between slopes of word recognition psychometric functions and homogeneity of the stimulus materials. J Am Acad Audiol. 2001;12(1):7-14.
11. Polspoel S, Holtrop FS, Bosman AJ, Kramer SE, Smits C. Measurement and optimisation of the perceptual equivalence of the Dutch consonant-vowel-consonant (CVC) word lists using synthetic speech and list pairs. Int J Audiol. 2025;64(1):35-42. [DOI: 10.1080/14992027.2024.2306186]
12. Cutler A, Weber A, Smits R, Cooper N. Patterns of English phoneme confusions by native and non-native listeners. J Acoust Soc Am. 2004;116(6):3668-78. [DOI: 10.1121/1.1810292]
13. Parikh G, Loizou PC. The influence of noise on vowel and consonant cues. J Acoust Soc Am. 2005;118(6):3874-88. [DOI: 10.1121/1.2118407]
14. Nishi K, Lewis DE, Hoover BM, Choi S, Stelmachowicz PG. Children's recognition of American English consonants in noise. J Acoust Soc Am. 2010;127(5):3177-88. [DOI: 10.1121/1.3377080]

Table 1- The mean recognition scores (%) of the three lists of SBMU -1 using whole-word scoring in participants (n = 22).

SNR (dB)	List 1		List 2		List 3		Freidman Chi-Square	P
	Mean	SD	Mean	SD	Mean	SD		
-5	28.61	14.55	33.34	15.74	33.42	14.37	6.52	0.038
0	65.09	9.81	73.09	8.96	69.09	9.27	10.50	0.005*
5	89.64	4.81	83.55	7.9	87.27	5.94	15.24	0.001*
10	92.09	4.39	95.18	4.03	95.82	2.38	11.18	0.004*
15	96.55	3.28	98.7	2.32	97.59	3.00	5.95	0.051*
20	98.64	1.89	99.27	1.32	97.82	1.94	8.49	0.014*

Table 2- The mean recognition scores (%) of the three lists of SBMU-1 using phoneme scoring in participants (n = 22).

SNR (dB)	List 1		List 2		List 3		Freidman Chi-Square	P
	Mean	SD	Mean	SD	Mean	SD		
-5	45.18	20.60	43.91	18.88	44.39	18.10	1.18	0.554
0	78.09	8.68	82.97	7.11	79.94	8.72	7.47	0.024
5	91.67	4.69	93.73	3.66	91.70	3.96	3.05	0.217
10	96.24	2.57	97.24	2.42	97.24	2.19	3.73	0.154
15	98.55	1.41	99.24	1.23	98.76	1.50	3.49	0.174
20	98.76	1.44	99.45	1.06	99.18	1.30	5.07	0.079

Table 3- The characteristics of the psychometric functions and the predicted SNR and slope values for the two scoring methods

Whole-word scoring				Phonemic scoring				List
Slope 20–80% (%/dB)	**SNR 80% (dB)**	**SNR 50% (dB)**	**SNR 20% (dB)**	**Slope 20–80% (%/dB)**	**SNR 80% (dB)**	**SNR 50% (dB)**	**SNR 20% (dB)**	
6.3	3.5	–2.4	–6.0	7.2	0.6	–4.5	–7.8	1
7.5	1.8	–3.1	–6.2	8.4	–0.2	–4.5	–7.3	2
6.6	2.7	–3.0	–6.5	7.6	0.3	–4.6	–7.6	3

Table 4- Count of consonants (manner of articulation) and vowels in the three lists of****-1 monosyllabic words

	Phonemes	List 1	List 2	List 3
Consonants	Stop	33	31	30
	Fricative	32	33	28
	Affricative	4	7	6
	Nasal	11	13	18
	Approximant	20	18	16
Vowels	æ	14	12	7
	ɑ	11	12	14
	e	2	4	4
	i	6	6	5
	o	5	3	7
	oʊ̯	0	2	4
	u	12	12	8

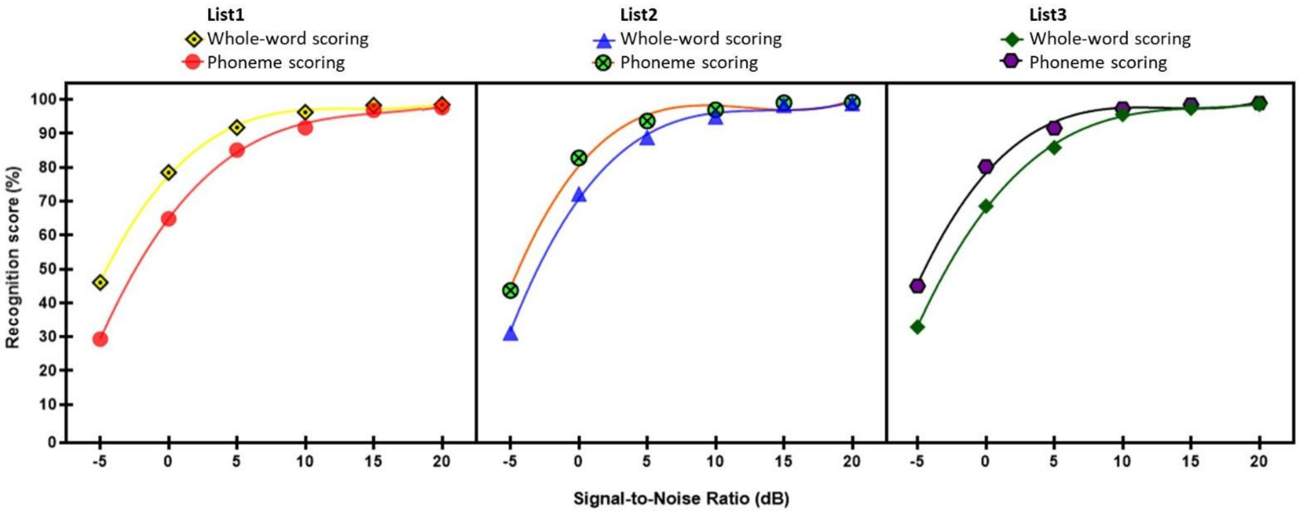

Figure ۱ -Psychometric function of mean recognition scores for the ×××-1 word lists with two scoring methods