

Speech-in-Noise Testing in Arabic: A Literature Review and Clinical Implications for Lebanese Arabic

Mostafa Darwish Akchair

SAERA. School of Advanced Education Research and Accreditation

ABSTRACT

Speech perception in noise is important for everyday communication and a key element used in hearing tests. Standard pure tone hearing tests do not fully assess functional hearing abilities in noisy environments and this is why speech in noise tests (SIN) are needed. Many speech-in-noise tests are created and validated in different languages, but there are a few standardized tests available in Arabic, particularly in local dialects - like Lebanese Arabic - remains limited. This study gives a review of earlier research that examines the pre-existing speech in noise tests, with a specific focus on Arabic speakers and dialect differences. A systematic search of major scientific databases was performed for studies published between the years of 1990 to 2025 that discussed the development, validation, and clinical use of speech in noise tests. Results indicate that most Arabic speech in noise tests use Modern Standard Arabic (MSA) or generalized dialect, which may not represent the ecological validity for Lebanese speakers. The review explains the strengths and weaknesses of current tools and points out the significant gap in validated SIN assessments tailored for Lebanese Arabic. These results show the need to develop dialect-based tests to improve clinical accuracy and research value. This thesis provides a foundation for future studies focused on creating and validating a Lebanese Arabic speech in noise test.

Keywords: *speech-in-noise, Lebanese Arabic, audiology, Arabic dialects, hearing assessments, test validation.*

INTRODUCTION

Effective speech communication often occurs in acoustically challenging environments, which makes the ability to understand speech in noise a critical auditory skill. People usually need to understand speech in the occurrence of background noise, competing talkers, reverberation, and intricate room acoustics; hence, facing auditory difficulties such as understanding speech in these situations is one of the main complaints that people with hearing impairment suffer from. Nonetheless, people with normal or near-normal pure-tone thresholds sometimes report similar complaints. As a result, basic hearing tests alone do not reveal the full extent of functional listening difficulties, which indicates why speech in noise testing is an important part of efficient hearing assessment.

In audiological sciences, speech-in-noise testing is crucial for evaluating speech recognition abilities and guiding clinicians in decision-making, particularly when fitting hearing aids or cochlear implants, and planning rehabilitation; however, many clinics still rely mainly on pure-tone audiometry, which does not reflect realistic listening conditions and, in turn, leads to underestimating a patient's true difficulties. Results from speech in noise tests have been shown to strongly predict how well a person communicates in real-world challenging listening environments (Grunditz & Magnusson, 2013). Speech understanding can be challenging when the speech signal is partially masked, distorted, or degraded by environmental factors such as background noise and reverberation; this makes accurate perception and linguistic processing even

more demanding (Charan et al., 2024). Additionally, SIN tests address different aspects of speech perception skills in distinct ways, and this factor alters test consistency and limits comparison between studies. Understanding speech in noise is considered a major obstacle for people with hearing impairments, especially for those with sensorineural hearing loss and central auditory processing disorders, who often have much poorer speech intelligibility in noisy environments compared to their audiometric hearing thresholds (Kılıç & Kara, 2023; Mehrez & Mohsen, 2023).

To tackle these challenges, several speech-in-noise tests have been developed, such as the Words in Noise test (WIN), Hearing in Noise Test (HINT), Quick Speech in Noise test (QuickSIN), and Bamford-Kowal-Bench Speech in Noise Test (BKB-SIN). These tests are commonly used in clinics and research settings and have been shown to provide valuable insight into functional hearing abilities, not only in quiet settings, but more importantly, in real-life complex listening situations. Most of these tests were first established in the English language and later translated into other Western languages. While some tests have been tested for clinical use and validated in non-Western groups, using them directly in different languages and linguistically diverse contexts and cultures can raise concerns regarding cultural and linguistic appropriateness. The reason behind this is that sentence structure, word familiarity, and listening habits may not match everyday speech, which consequently affects the accuracy of the results (Sharma et al., 2016; Skoe & Karayanidi, 2019). Therefore, such tests do not evaluate real communication abilities among diverse populations, which states a

well-defined need to establish better language- and culture-specific tools.

In Arabic-speaking countries, testing speech perception in noise can be considered complicated due to the diglossic nature and substantial variation in the Arabic language among regional dialects. Modern Standard Arabic (MSA) is mainly used in formal speech and writing, while daily communication happens in local dialects that can vary remarkably between one region and the other. Lebanese Arabic, which is a Levantine dialect, has distinctive phonological, lexical, and prosodic features that differentiate it from other Arabic dialects. The Arabic language has 28 consonant sounds and six phonemes. In many modern dialects, the vowel inventory may expand to eight phonemes, including some phonemes that do not exist in other languages, and these phoneme features interact with dialect-specific patterns of stress, intonation, and rhythm (Alqarhi, 2019).

It is important to note that there is clear dialectic variation within Lebanon, specifically between urban and rural areas. In urban contexts, speakers often switch between three languages (Arabic, English, and French) during daily life conversation, which leads to unique stress patterns, intonation contours, and prosodic features. These features strongly affect how well speech is perceived, especially in noisy environments where listeners rely heavily on prosodic and contextual cues. Overlooking such features leads to unrealistic test results that mismatch the patient's real-world listening skills. Therefore, it is essential to utilize test materials that match the local dialect when assessing speech perception in noise.

Despite the spread of awareness towards audiological assessments in the Arab region and in Lebanon in particular, a huge gap still exists in designing standardized SIN tests for Lebanese Arabic. Current Arabic speech-in-noise tests have been developed using Moroccan, Iraqi, Saudi, Egyptian, Jordanian, and Palestinian Arabic dialects (Alokaily et al., 2022). While these tests are useful in their own countries, they are thoroughly adequate for Lebanese speakers. Linguistic mismatch between test materials and the listener's native dialect leads to poor results due to unfamiliar vocabulary, phonology, or prosody, rather than real auditory deficits (Bent & Bradlow, 2003). Accordingly, these tests do not extensively reflect real-life listening abilities for Lebanese Arabic speakers, which results in the reduction of ecological validity.

The main objective of this study is to review studies conducted in the Arab region that present speech-in-noise testing in Arabic for Arabic-speaking subjects. This review aims to help guide audiologists in Lebanon in establishing and validating a Lebanese Arabic speech-in-noise test that better serves the Lebanese dialect and culture. Moreover, the introduction delves into the formulation of research objectives, delineating the specific aims that guide the research trajectory. The theoretical framework, a conceptual scaffold upon which the study rests, is meticulously articulated to elucidate the intellectual underpinnings shaping the inquiry. Importantly, the methodological approach is expounded upon, providing a transparent exposition of the strategies employed in data collection and analysis.

REVIEW OF LITERATURE

Research on SIN testing in Arabic-speaking populations has expanded considerably over the past decade, with increasing efforts to adapt, validate, and standardize SIN tools for clinical and research use. These studies differ in test design, linguistic approach, and clinical application, providing valuable insight into the strengths and limitations of existing Arabic SIN assessments.

One of the most influential contributions is the development of an Arabic HINT-type test using sentence-based adaptive methodology. Essawy et al. evaluated this test in a large sample of adults with normal hearing ($n=150$) and a smaller group with sensorineural hearing loss ($n=30$). SRTs were measured in quiet and noise presented from different spatial locations. The results demonstrated good reliability, clinical efficiency, and sensitivity to spatial noise configuration, supporting the usefulness of this test for hearing-aid evaluation and outcome measurements. However, the sentence materials were based primarily on Egyptian Arabic and formal linguistic structures, which may reduce familiarity and ecological validity for listeners from other dialect regions, including Lebanon.

The clinical applicability of Arabic HINT-based testing was further examined by Kamal et al. (2023), who used the test to compare fixed and adaptive directional microphone technologies in hearing aids. Their findings showed significantly better SNR thresholds with adaptive directional microphones, confirming the sensitivity of sentence-based SIN testing to amplification technology. Nevertheless, the dialectal specificity of the test materials remains a

limitation when results are generalized to non-Egyptian Arabic-speaking population.

Adaptations of the QuickSIN test have been reported by Elrifayy et al. (2021) that was developed and standardized in Arabic that preserved the original test structure, including multi-talker babble and stepwise SNR changes, allowing rapid clinical estimation of SNR loss. While the test was shown to be time-efficient and clinically practical, its reliance on fixed sentence lists increases vulnerability to linguistic bias when sentence familiarity differs across dialects. Sultan et al. (2020) compared Arabic QuickSIN outcomes with Arabic HINT results in adults with SNHL and emphasized that QuickSIN's speed advantage must be balanced against potential reductions in sensitivity and ecological validity, particularly when sentence materials are not dialect-matched.

Word-based SIN testing has been explored primarily in pediatric populations. Abdel Rahman (2018) developed an Arabic WIN test for Egyptian children ($n=152$) and provided normative data demonstrating the test's ability to distinguish children with normal development from those with CAPD. This study highlights the importance of age-appropriate and familiar vocabulary in pediatric SIN testing. From a Lebanese perspective, however, the use of Egyptian dialect materials limits generalizability, underscoring the lack of pediatric SIN tools that reflect everyday spoken Lebanese Arabic.

Matric sentence test have gained attention due to their high level of standardization and cross-language comparability. Asiri et al. (2025) examined the test-retest reliability of an Arabic speech matric test in listeners with normal hearing, hearing-aid users, and CI

users. the study reported minimal learning effects despite this advantage, matrix sentences are typically constructed using standardized lexical items that resemble MSA more than spontaneous spoken dialects, which may reduce ecological validity for Lebanese Arabic listeners.

Digits-in-Noise tests have been increasingly applied for hearing screening and remote assessment in Arabic-speaking populations. In which Arabic DIN tests identifies hearing difficulties across large populations although results may be affected by device type, listening environments, and language - related factors (Masalski & Morawski, 2025b). such simplicity also limits the test's ability to reflect real-world communications demands that involve sentence-level processing in noise (Schimmel et al., 2024)

Speech-in-noise tests are primarily helpful in indicating hidden hearing difficulties, such as issues associated with cochlear synaptopathy, auditory processing disorders, or age-related changes in temporal and cognitive processing (Guo et al., 2025). In such situations, people may perform well in hearing tests in quiet but still show significant deficits in noisy environments. Therefore, speech-in-noise testing better measures the auditory performance than standard PTA tests.

In clinical settings, speech-in-noise testing is very crucial in the assessment and management of hearing aids and cochlear implants. The test outcome in noise better evaluates the extent of the patient's satisfaction with the amplification device than other speech tests performed in quiet. In addition, speech-in-noise tests are often used to measure the functionality and application of modern hearing aid features, such as directional microphones and noise-reduction

systems, to show improved SNR under controlled test conditions (Torkildsen et al., 2019).

Speech in noise tests extend beyond assessment and treatment to include counseling and clinical decision-making. Validating patients' complaints through a clear representation of outcomes helps build trust and facilitate communication between the patient and the clinician. Moreover, it encourages realistic expectations towards the extent of benefit that the amplification device can provide. Speech-in-noise test results guide personalized treatment plans and support-evidence based choices about hearing aids, assistive listening devices, and useful communication tools for everyday life (Davidson et al., 2022).

This testing strategy plays a crucial role in modern hearing assessment as it provides a clearer picture of a person's performance in noisy listening situations, thereby including these tests in routine hearing evaluations helps clinicians accurately diagnose, improve outcome evaluation, and provide more targeted hearing care towards an individual's needs.

Speech recognition in noise is a difficult auditory task that depends on audibility, temporal-spectral resolution, linguistic processing, and higher-level cognitive resources. In standard hearing tests, pure-tone audiometry (PTA) measures hearing sensitivity, but it does not clearly imply the communication struggles that a person may face in real-life noisy situations. Therefore, speech-in-noise (SIN) tests are recommended as additional tools to quantify everyday listening problems and assist with amplification device fittings (McCreery et al., 2019).

There are two main types of SIN tests commonly used in clinics. One that includes sentence-based adaptive tests that measure the speech reception threshold (SRT), also called SNR-50, which is the signal-to-noise ratio needed for a person to understand 50% of speech correctly. The other type includes fixed-list tests, which are often word-based and measure the number of words understood correctly at fixed noise levels. Sentence-based tests, such as HINT-style tests, are faster and present clear numerical values (dB SNR) while word-based tests are specifically helpful for children in addition to studying auditory processing when carefully controlled for lexical familiarity and developmental language level (Billings et al., 2024).

Several international SIN tests are commonly used as methodological models when new tests are translated or adopted in other languages. HINT is an adaptive sentence-based test that estimates the speech reception threshold by adjusting either the speech or noise level in quiet and noisy environments. The QuickSIN is a fast clinical tool that measures SNR loss using fixed sentence lists presented with multi-talker babble at gradually increasing noise levels. The WIN test uses monosyllabic words in competing noise to assess word recognition and auditory closure abilities (Billings et al., 2024). Matrix sentence tests rely on sentences built from fixed word structure, allowing strong standardization, efficient adaptive testing, and comparison across languages (Kollmeier et al., 2015b). DIN test uses triple-digit numbers in noise, and it is widely applied for screening and remote assessment due to ease of administration and minimal language skills requirement (Masalski & Morawski, 2025a). Together, these test types illustrate an

important balance between ecological validity, meaning natural everyday language use, and the need for standardization and cross-language comparability in SIN assessments.

Arabic has a diglossia structure, meaning that Modern Standard Arabic (MSA) is mainly used in formal situations, while people use local spoken dialects in everyday life. This creates a challenge when developing speech-in-noise tests since MSA helps develop standardized tests across different countries; however, it can reduce ecological validity in everyday listening contexts (Zokoll et al., 2025). Research on Arabic speech-in-noise testing has increased in the last ten years, with many test development and validation studies coming from Egypt and the Gulf countries. Despite this progress, Arabic speech-in-noise tools are still limited by gaps across different dialects, age groups, and clinical populations, which shows that more balanced and inclusive test development needs to be implemented (Alokaily et al., 2022).

An important contribution to Arabic SIN testing is the creation and standardization of an Arabic HINT-type test that uses 28 balanced lists of everyday sentences and an adaptive testing method. In a large group of adults with normal hearing (n=150) and a smaller clinical group with mild to moderate sensorineural hearing loss (n=30), Essawy et al. measured speech reception thresholds in both quiet and SNR conditions with noise coming from different directions. They found that the test was reliable, time-efficient, and suitable for clinical use. They also indicated that using noise from multiple directions is useful for hearing aid validation. As a result, this test provides a strong

methodological model for sentence-based Arabic SIN assessment, particularly for hearing-aid research and fitting (Essawy, 2019).

The clinical value of this approach was further shown by Kamal et al. (2023), who used the Arabic HINT to compare different hearing aid microphone technologies, including fixed and adaptive directional microphones, which usually produce better (lower) SNR thresholds, supporting the usefulness of Arabic sentence-in-noise testing for measuring hearing aid benefit (Kamal et al., 2023). Nevertheless, if sentence materials are based on Egyptian dialect features or more formal MSA-like structures, Lebanese listeners may face unfamiliar words or speech patterns. This dialect mismatch could affect performance and reduce the test effectiveness for Lebanese speakers, highlighting the need for dialect-appropriate materials.

The QuickSIN approach has also been adapted for use in Arabic. Elrifay et al. (2021) described the development and standardization of an Arabic QuickSIN designed to quickly measure signal-to-noise ratio loss in clinical settings, while keeping the original test structure of multi-talker babble and stepwise changes in SNR to allow time-efficient assessment (Elrifay et al., 2021).

Sultan et al. (2020) later compared the Arabic QuickSIN with the Arabic Hearing in Noise Test in adults with sensorineural hearing loss while focusing on their clinical usefulness and sensitivity. Such comparisons are important because QuickSIN is often preferred for its speed and simplicity; whereas HINT-type tests provide adaptive threshold estimation and allow more flexibility in noise direction. Moreover,

QuickSIN sentence materials must be carefully controlled for linguistic difficulty because unfamiliar dialect features can artificially increase measured SNR loss and confuse clinical interpretation (Sultan et al., 2020). This concern is particularly relevant in Lebanese Arabic, where differences in vocabulary and sentence structure from Egyptian-based or MSA-influenced materials may reduce ecological validity and highlight the need for dialect-specific adaptations.

Rahman (2018) developed an Arabic WIN test for Egyptian children aged 5-12 years (n=152), whereby he provided normative data and showed that this test can clearly distinguish between normally developing children and those with central auditory processing disorders. Furthermore, this study emphasizes the importance of choosing words that are familiar and appropriate for a child's age and language level and proves that SIN tests can be successfully implemented in Arabic language (Abdel Rahman, 2018). From a Lebanese Arabic perspective, this work also highlights a major limitation, as pediatric SIN materials are supposed to rely heavily on everyday spoken language to reflect how children actually communicate. Lebanese Arabic pediatric SIN tests are still very limited, which creates a significant gap in early diagnosis and timely intervention for Lebanese children with listening difficulties in noise.

Matrix SIN tests have become popular because they use fixed sentence structure and allow results to be easily compared across different languages. Recent research has shown that Arabic matrix tests are reliable and practical for clinical use. For example, Asiri et al. (2025) studied the test-retest reliability of the Arabic Speech Matrix Test

in people with normal hearing as well as hearing aid and cochlear implant users, and they found moderate to excellent reliability with only minimal learning effects, as these findings support the usefulness of matrix tests for both research and clinical assessment (Asiri et al., 2025). Nevertheless, a key limitation remains, as matrix sentences often rely on standardized words and sentence patterns that are closer to MSA. This does not accurately reflect everyday spoken Lebanese Arabic, although it improves consistency and cross-language comparison. Hence, without careful adaptation, this mismatch can reduce ecological validity and limit the accuracy of the representation of real-life listening for Lebanese speakers.

Arabic DIN tests are being used more often for hearing screening and remote assessments. Recent validation studies, including browser-based DIN tests, reveal that Arabic DIN can quickly and efficiently measure hearing thresholds, although results can be affected by the type of device used, the listening environment, and language-related factors (Masalski & Morawski, 2025b). On one side, the DIN test uses numbers instead of full sentences, which makes it they reduce language difficulty and an easy test to be used with large groups of people and reduces the possibility of language difficulty. On the other side, this simplicity is also a limitation, as DIN tests may not reflect real-life listening challenges that involve understanding and comprehending full sentences in everyday noisy situations (Schimmel et al., 2024a).

Across Arabic SIN research, a clear and repeated limitation is the weak inclusion of dialect differences in test material design and result interpretation. Arabic is diglossic and

includes many regional dialects; hence, using Modern Standard Arabic or non-local spoken varieties can introduce systematic bias into SIN testing. For Lebanese Arabic in particular, phonological distinctions such as vowel quality patterns, prosody, and lexicon differentiate it from Egyptian and Gulf dialects (Bergstrand Othman, 2025). These linguistic differences can affect speech understanding in noise at both the acoustical level (how sounds are perceived) and the cognitive level (how words are recognized and understood). As a result, poorer performance may reflect dialect mismatch rather than true hearing difficulty. A review by Rawas (2023) highlights this issue, stressing the lack of validated, dialect-representative Arabic SIN materials and confirming that accent and dialect mismatch can negatively influence test outcomes and reduce clinical usefulness (Rawas, 2023).

Overall, the reviewed literature demonstrated that while several Arabic SIN tests are reliable and clinically valuable within their target populations, none adequately address the linguistic and ecological needs of Lebanese Arabic speakers. This gap reinforced the need for the development and validation of a Lebanese Arabic SIN test that balances standardization with dialectal authenticity to ensure accurate assessment of real-world listening abilities.

METHOD

Aim of the study

The primary aim of this study is to systematically review and critically appraise the existing body of literature on SIN testing in Arabic-speaking populations, with particular emphasis on the availability,

validity, and clinical relevance of such assessments for Lebanese Arabic speakers. SIN testing is widely acknowledged as a core component of contemporary audiological evaluation because it captures functional hearing abilities under ecologically valid listening conditions that are not reflected by PTA alone. Despite this consensus, the extent of currently available SIN tests that adequately account for the linguistic heterogeneity of Arabic remains unclear, particularly dialectal variation and diglossia, which are known to influence speech perception in noise. This review critically examines the methodological characteristics of Arabic SIN studies, including stimulus design, participant selection, validation procedures, dialect choice, reliability metrics, and normative frameworks, to evaluate their scientific rigor and clinical applicability.

Inclusion criteria

The review synthesizes evidence from peer-reviewed journal articles, conference proceedings, and academic theses that address SIN test development, validation, or application in the Arabic-speaking population, with particular attention to studies reporting psychometric properties such as test-retest reliability, hearing loss sensitivity, and normative data. The literature demonstrates growing activity in Arabic SIN research, especially in Egypt and the Gulf region; however, critical gaps persist, including limited dialectal representativeness, underrepresentation of diverse clinical populations, and inconsistent consideration of linguistic load as a confounding factor.

A further objective of this review is to explicitly identify these gaps, mostly the

absence of validated SIN assessments tailored for Lebanese Arabic, and to argue that reliance on non-local dialects or MSA-based materials may compromise ecological validity and clinical interpretability. By highlighting these limitations, this study provides an evidence-based rationale for future research aimed at developing and validating linguistically and culturally appropriate SIN tools for Lebanese Arabic speakers, therefore supporting more accurate diagnosis, outcome evaluation, and patient-centered audiological care.

Exclusion criteria

The exclusion criteria for this review were defined to ensure methodological rigor. Studies that were not available in full text were excluded because they do not allow comprehensive methodological evaluation; research unrelated to audiological assessment was excluded to maintain clinical focus, such that studies that examined speech-in-noise or related phenomena exclusively in non-Arabic languages were excluded as they cannot support Arabic-specific conditions. Moreover, narrative reviews, opinion pieces, and other publications that lack original data were excluded because they do not contribute primary evidence on test development, validation, or performance.

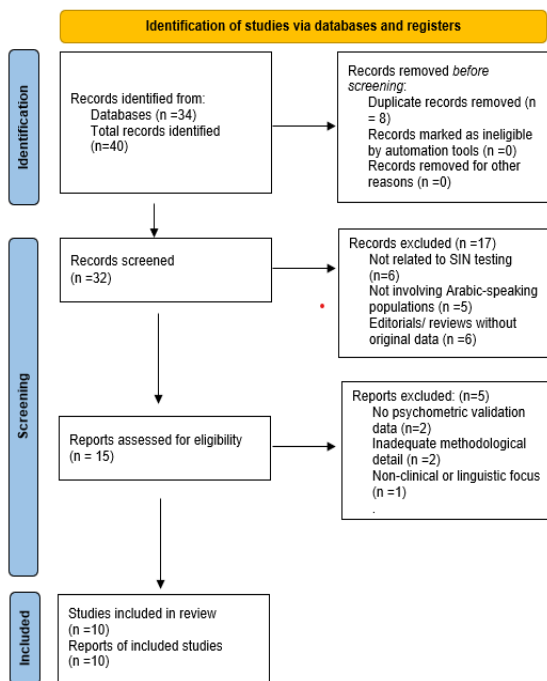
Procedure

A comprehensive and systematic literature search was conducted using major electronic databases, some of which are PubMed, ScienceDirect, Web of Science and Google Scholar. An initial screening of titles and abstracts was performed to assess relevance to SIN assessments in Arabic-speaking populations, followed by a full-text review of

studies meeting the inclusion criteria. To ensure completeness, the reference lists of all included articles were manually examined to identify additional relevant publications not captured through database searching.

Figure 1.

PRISMA Flow Diagram



RESULTS

The synthesis of the reviewed literature revealed many consistent trends in the development, validation, and use of SIN tests for Arabic-speaking populations. Findings were organized thematically to highlight shared methodological approaches, psychometric outcomes, and linguistic considerations across studies.

A total of ten studies met the inclusion criteria and were analyzed in this review. These studies reported outcomes related to SIN performance in Arabic-speaking

populations using sentence-based, matrix, word-based, and digits-in-noise test formats. Reported findings included sample characteristics, normative values, speech reception thresholds, SNR loss, and reliability measures.

Essawy et al. (2019) evaluated an Arabic HINT-type test in 150 adults with normal hearing and 30 adults with mild-to-moderate sensorineural hearing loss. Mean speech reception thresholds for normal-hearing participants ranged approximately -7 to -9 dB SNR depending on noise direction, while listeners with hearing loss required significantly higher SNRs. The study reported good test-retest reliability and demonstrated sensitivity to spatial noise configurations.

Kamal et al. (2023) applied the Arabic HINT in adults using hearing aids to compare microphone technologies. Participants showed significantly lower (better) SRTs when using directional microphones compared to fixed directional or omnidirectional settings, confirming the test’s sensitivity to amplification features.

Elrifayy et al. (2021) standardized an Arabic version of the QuickSIN test in adults with normal hearing. Normative SNR-loss values were reported, with most normal-hearing participants demonstrating SNR losses within the expected clinical range (0-2 dB). The test demonstrated acceptable internal consistency and was time-efficient for clinical use.

Sultan et al. (2020) compared the Arabic QuickSIN and Arabic HINT outcomes in adults with SNHL. results showed that QuickSIN produced higher SNR-loss values than adaptive HINT SRTs, with variability attributed to sentence familiarity and fixed

presentation levels. both tests successfully differentiated normal-hearing and hearing-impaired participants.

Abdel Rahman (2018) developed an Arabic Words-in-noise test for Egyptian children aged 5-12 years (n=152). Age-specific normative data were established, showing gradual improvement in word recognition in noise with increasing age. The test successfully distinguished children with central auditory processing disorders from typically developing peers.

Asiri et al. (2025) investigated the Arabic speech matrix test in adults with normal hearing, hearing-aid users, and CI users. Test-retest reliability ranged from moderate to excellent across groups, with ICC exceeding 0.80 and minimal learning effects were observed between sessions.

Masalski and Morawski (2025a) reported results from Arabic DIN testing conducted in controlled and remote settings. Normal-hearing participants obtained lower digit recognition thresholds than listeners with hearing loss, and results were consistent across test platforms, although increased variability was noted in unsupervised environments.

Across the included studies, all Arabic SIN tests demonstrated the ability to differentiate between normal-hearing and hearing-impaired listeners. Sentence-based adaptive tests consistently produced quantifiable SRT values, WIN tests provided age-based pediatric norms, matrix tests showed strong reliability coefficients, and DIN tests yielded efficient screening thresholds.

None of the ten studies reported normative data, validation outcomes, or reliability measures for a speech-in-noise test developed specifically in Lebanese Arabic,

and no Lebanese-specific adult or pediatric norms were identified.

Overall, the evidence suggests that Arabic SIN tests are possible to use, reliable, and helpful in clinical settings; however, when these studies are examined more closely with respect to dialect differences and their suitability for Lebanese Arabic speakers (Table 1.), important limitations appear. These issues reduce the applicability of results for Lebanese listeners and limit the generalization of the findings.

DISCUSSION

The studies reviewed show clear progress in developing and using speech-in-noise (SIN) tests for Arabic-speaking populations.

Essawy et al. (2019) is one of the strongest studies on Arabic speech-in-noise testing. The study included 150 adults with normal hearing as well as a group (n=30) with sensorineural hearing loss, which makes the results reliable and allows clear comparison between groups. The use of adaptive sentence testing and different noise directions can imitate real-life listening situations and present useful information for clinical purposes such as hearing aid evaluation. On the other hand, this study showed a mismatch of the used dialect with the targeted population, whereby the Egyptian accent was applied for Lebanese speakers (Essawy, 2019). Dialect differences are known to affect speech understanding in noise; hence, the lack of testing with Lebanese Arabic limits how well these results can be applied to Lebanese listeners.

The Arabic QuickSIN test adaptation by Elrifayy et al. (2021) is strengthened by its ease of use which makes it suitable for

routine clinical practice; however, the study shows limited detail on participant characteristics and norms, reducing its overall strength. QuickSIN uses a small set of fixed sentences, influencing dialect familiarity (Elrifayy et al., 2021). As a result, this test indicates that quick SNR-loss measurement in Arabic is possible; nonetheless, its accuracy for Lebanese Arabic speakers remains limited.

Sultan et al. (2020) adds useful evidence by directly comparing Arabic QuickSIN and Arabic HINT in adults with normal hearing and those with hearing loss. This direct comparison helps show differences in the sensitivity and clinical usefulness of each test, suggesting that adaptive sentence-based tests can better reflect real-life listening than fixed-list test (Sultan et al., 2020). Nevertheless, the study used a small sample size and non-Lebanese Arabic materials, the fact that limit the applicability of results on Lebanese speakers. Therefore, the findings are best viewed as relevant to one Arabic dialect rather than all Arabic-speaking populations.

Rahman's (2018) study provides strong age-based norms and is one of the few Arabic speech-in-noise tests designed for children. The large sample size and the inclusion of central auditory processing disorder strengthen the study's evidence; however, word-based tests depend heavily on lexical familiarity and phonological structure, so they are likely to be affected by dialect differences (Abdel Rahman, 2018). Therefore, differences in vocabulary and pronunciation in Lebanese Arabic may affect children's performance, limiting how well these norms apply to Lebanese speakers.

Kamal et al. (2023) provide useful clinical evidence by showing that Arabic HINT

results are sensitive to hearing aid microphone directionality, supporting the use of Arabic SIN tests for outcome measures, not only as diagnostic tools; however, the study assumes that all Arabic dialects are linguistically equivalent and does not consider the effect of dialect mismatch on results (Kamal et al., 2023). Accordingly, the findings do not fully address validity for Lebanese speakers although they are still considered to be clinically helpful.

Rawas (2023) offers a strong theoretical and methodological contribution by clearly discussing the limitations of current Arabic SIN tools, especially related to dialect differences (Rawas, 2023). Although the work is not specific to Lebanese Arabic and lacks Lebanese norms, it strengthens the evidence base through its comprehensive analysis and provides important guidance for future test development.

Overall, the evidence supporting Arabic SIN testing is methodologically strong but limited by context. The reviewed studies show good reliability, clinical usefulness, and practicality, but their strength decreases when applied to speakers of different Arabic dialects. This limitation supports the main argument of this thesis: that SIN tests must be developed and validated for specific dialects to reflect real-life listening accurately. For Lebanese Arabic speakers, existing studies represent an important milestone, but not the endpoint, highlighting the need for dedicated test development based on authentic language use and strong validation methods.

Future Directions

Future research should build onto this review by focusing on creating and validating SIN

tests that match specific Arabic dialects; Lebanese Arabic in particular. Knowing that dialect differences affect speech understanding in noise, future studies should carefully design test sentences that reflect spoken Lebanese Arabic in daily life by including phonological, lexical, and prosodic features. These materials should be tested in advance to ensure familiarity and naturalness for listeners. Large studies are also needed to develop clear norms for different ages and hearing levels to interpret clinical results accurately. Moreover, researchers should also assess the reliability of tests with respect to its duration and the accuracy of detecting hearing loss and benefits from amplification devices. Comparing results across Arabic dialects may also help explain how language differences affect performance. Using digital or remote testing methods further improves access and practicality. Therefore, all these steps should be implemented to create efficient and effective SIN assessments for Lebanese Arabic speakers and other underserved Arabic speaking groups.

CONCLUSION

This thesis reviewed the existing evidence on SIN testing in Arabic-speaking populations, focusing on how well current tools apply to Lebanese Arabic speakers. The review shows that while significant progress has been made in adapting and validating SIN tests in Arabic, important linguistic and methodological limitations still exist. These limitations reduce the accuracy and clinical usefulness of existing tools for Lebanese speakers.

The studies reviewed demonstrate that Arabic SIN tests are feasible, reliable, and informative. Sentence-based adaptive tests,

such as Arabic HINT-type measures, are sensitive to hearing loss and hearing-aid benefits, while rapid tests like Arabic QuickSIN provide efficient estimations of SNR loss. Pediatric word-in-noise tests further emphasize the value of SIN assessment across age groups. This evidence altogether confirms that SIN testing provides insights beyond pure-tone audiometry into functional hearing ability.

The most consistent limitations across studies is the lack of dialect-specific validation. Most tests use Egyptian Arabic or MSA materials without considering regional dialect differences. Given that lexical familiarity, phonology, and prosody strongly affect SIN perception, this mismatch can bias results and limit the applicability of normative data for Lebanese Arabic speakers, reducing ecological validity.

Comparisons of test formats indicate that no single SIN test serves all purposes. Sentence-based tests are highly realistic but sensitive to dialect differences, while WIN and DIN based tests reduce linguistic demands but are less reflective of everyday communication. These findings highlight the importance of matching test design to both clinical and linguistic goals.

As a conclusion, the current Arabic SIN tests evidence provides a strong methodological foundation; however, it does not entirely meet the needs of Lebanese Arabic speakers. This thesis supports the development and validation of a Lebanese Arabic SIN test based on authentic language use and rigorous psychometric standards. Such tool is essential for improving clinical assessment, strengthening research validity, and ensuring equitable audiological care for Lebanese populations.

REFERENCES

- Abdel Rahman, T. T. (2018). Development and standardization of Arabic words in noise test in Egyptian children. *International Journal of Pediatric Otorhinolaryngology*, *108*, 1–7. <https://doi.org/10.1016/j.ijporl.2017.11.019>
- Alokaily, A. O., Alqabbani, A. F., Aleid, A., & Alhussaini, K. (2022). Toward Accessible Hearing Care: The Development of a Versatile Arabic Word-in-Noise Screening Tool: A Pilot Study. *Applied Sciences*, *12*(23), 12459. <https://doi.org/10.3390/app122312459>
- Alqarhi, A. (2019). Arabic Phonology. *English Linguistics Research*, *8*(4), 9. <https://doi.org/10.5430/elr.v8n4p9>
- Asiri, A., Almuhawwas, F., Alrushaydan, D., Aljabr, M., Mesallam, T. A., & Yousef, M. (2025). Reliability Analysis of the Arabic Speech Matrix Test. *Laryngoscope Investigative Otolaryngology*, *10*(2), e70138. <https://doi.org/10.1002/liv.2.70138>
- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, *114*(3), 1600–1610. <https://doi.org/10.1121/1.1603234>
- Bergstrand Othman, L. (2025). Arabic diglossia: Advocating for a non-deficit model in comparative analysis of reading and language acquisition. *Frontiers in Education*, *10*, 1518728. <https://doi.org/10.3389/educ.2025.1518728>
- Billings, C. J., Olsen, T. M., Charney, L., Madsen, B. M., & Holmes, C. E. (2024). Speech-in-Noise Testing: An Introduction for Audiologists. *Seminars in Hearing*, *45*(01), 055–082. <https://doi.org/10.1055/s-0043-1770155>
- Charan, G., Sivateja, G., Karthik Sarma, M., Sabyasachi Kumar. (2024). *Unveiling the Challenges of Speech Recognition in Noisy Environments: A Comprehensive Review of Issues and Solutions*. Unpublished. <https://doi.org/10.13140/RG.2.2.24231.76966>
- Davidson, A., Marrone, N., & Souza, P. (2022). Hearing Aid Technology Settings and Speech-in-Noise Difficulties. *American Journal of Audiology*, *31*(1), 21–31. https://doi.org/10.1044/2021_AJA-21-00176
- Elrifayy, M., Emara, A., Mahalawy, T., & EI-Gharib, A. (2021). Development and Standardization of Arabic Version of Quick Speech in Noise Test. *Egyptian Journal of Ear, Nose, Throat and Allied Sciences*, *22*(22), 1–9. <https://doi.org/10.21608/ejentas.2021.35054.1235>
- Essawy, W. M. (2019). Application of Arabic hearing in noise test on subjects with sensorineural hearing loss. *International Journal of Otorhinolaryngology and Head and Neck Surgery*, *5*(6), 1635.

- <https://doi.org/10.18203/issn.2454-5929.ijohns20194940>
- Grunditz, M., & Magnusson, L. (2013). Validation of a speech-in-noise test used for verification of hearing aid fitting. *Hearing Balance and Communication, 11*(2), 64–71. <https://doi.org/10.3109/21695717.2013.782135>
- Guo, X., Benzaquén, E., Holmes, E., Berger, J. I., Brühl, I. C., Sedley, W., Rushton, S. P., & Griffiths, T. D. (2025). Predicting speech-in-noise ability with static and dynamic auditory figure-ground analysis using structural equation modelling. *Proceedings of the Royal Society B: Biological Sciences, 292*(2042), 20242503. <https://doi.org/10.1098/rspb.2024.2503>
- Kamal, W. A., Essawy, W. M., Emara, A. A., & El-Gharib, A. M. (2023). Using the hearing in noise test (HINT) in the assessment of the performance of directional microphone. *The Egyptian Journal of Otolaryngology, 39*(1), 23. <https://doi.org/10.1186/s43163-023-00390-7>
- Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 116*(4), 2395–2405. <https://doi.org/10.1121/1.1784440>
- Kılıç, M., & Kara, E. (2023). Investigation of hearing aid users' speech understanding in noise and their spectral-temporal resolution skills. *Journal of Otology, 18*(3), 146–151. <https://doi.org/10.1016/j.joto.2023.04.001>
- Kollmeier, B., Warzybok, A., Hochmuth, S., Zokoll, M. A., Uslar, V., Brand, T., & Wagener, K. C. (2015a). The multilingual matrix test: Principles, applications, and comparison across languages: A review. *International Journal of Audiology, 54*(sup2), 3–16. <https://doi.org/10.3109/14992027.2015.1020971>
- Kollmeier, B., Warzybok, A., Hochmuth, S., Zokoll, M. A., Uslar, V., Brand, T., & Wagener, K. C. (2015b). The multilingual matrix test: Principles, applications, and comparison across languages: A review. *International Journal of Audiology, 54*(sup2), 3–16. <https://doi.org/10.3109/14992027.2015.1020971>
- Masalski, M., & Morawski, K. (2025a). The multilingual digits-in-noise (DIN) test: Development and evaluation. *International Journal of Audiology, 64*(11), 1193–1203. <https://doi.org/10.1080/14992027.2024.2397068>
- Masalski, M., & Morawski, K. (2025b). The multilingual digits-in-noise (DIN) test: Development and evaluation. *International Journal of Audiology, 64*(11), 1193–1203. <https://doi.org/10.1080/14992027.2024.2397068>

- McCreery, R. W., Walker, E. A., Spratford, M., Lewis, D., & Brennan, M. (2019). Auditory, Cognitive, and Linguistic Factors Predict Speech Recognition in Adverse Listening Conditions for Children With Hearing Loss. *Frontiers in Neuroscience, 13*, 1093. <https://doi.org/10.3389/fnins.2019.01093>
- Mehrez, R. N., & Mohsen, S. M. (2023). Study of the Effect of Minimal Sensorineural Hearing Loss on the Central Auditory Processing in Children: A Pilot Study. *Indian Journal of Otolaryngology and Head and Neck Surgery: Official Publication of the Association of Otolaryngologists of India, 75*(4), 2781–2785. <https://doi.org/10.1007/s12070-023-03862-7>
- Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *The Journal of the Acoustical Society of America, 95*(2), 1085–1099. <https://doi.org/10.1121/1.408469>
- Rawas, I. O. (2023). *Dataset supporting the University of Southampton Doctoral Thesis “Developing an Arabic speech in noise test as a measure of auditory fitness for duty” by Iman Osamah Rawas*. [Dataset]. University of Southampton. <https://doi.org/10.5258/SOTON/D2626>
- Schimmel, C., Cormier, K., Manchaiah, V., Swanepoel, D. W., & Sharma, A. (2024). Digits-in-Noise Test as an Assessment Tool for Hearing Loss and Hearing Aids. *Audiology Research, 14*(2), 342–358. <https://doi.org/10.3390/audiolres14020030>
- Sharma, S., Tripathy, R., & Saxena, U. (2016). Critical appraisal of speech in noise tests: A systematic review and survey. *International Journal of Research in Medical Sciences, 5*(1), 13. <https://doi.org/10.18203/2320-6012.ijrms20164525>
- Skoe, E., & Karayanidi, K. (2019). Bilingualism and Speech Understanding in Noise: Auditory and Linguistic Factors. *Journal of the American Academy of Audiology, 30*(2), 115–130. <https://doi.org/10.3766/jaaa.17082>
- Sultan, O., Elmahallawy, T., Kolkaila, E., & Lasheen, R. (2020). Comparison Between Quick Speech in Noise Test (QuickSIN test) and Hearing in Noise Test (HINT) in Adults with Sensorineural Hearing Loss. *Egyptian Journal of Ear, Nose, Throat and Allied Sciences, 0*(0), 0–0. <https://doi.org/10.21608/ejentas.2020.28080.1195>
- Torkildsen, J. V. K., Hitchins, A., Myhrum, M., & Wie, O. B. (2019). Speech-in-Noise Perception in Children With Cochlear Implants, Hearing Aids, Developmental Language Disorder and Typical Development: The Effects of Linguistic and Cognitive Abilities. *Frontiers in Psychology, 10*, 2530.

- <https://doi.org/10.3389/fpsyg.2019.02530>
- Turudu, S., Koelewijn, T., Gaudrain, E., Illan, G. A., Rachman, L., & Başkent, D. (2022). *An open generic specification for the Digits-in-Noise test*.
<https://doi.org/10.13140/RG.2.2.33645.46564/1>
- Vermiglio, A. J., Soli, S. D., Freed, D. J., & Fang, X. (2020). The Effect of Stimulus Audibility on the Relationship between Pure-Tone Average and Speech Recognition in Noise Ability. *Journal of the American Academy of Audiology*, 31(3), 224–232.
<https://doi.org/10.3766/jaaa.19031>
- Wilson, R. H., & Watts, K. L. (2012). The Words-in-Noise Test (WIN), List 3: A Practice List. *Journal of the American Academy of Audiology*, 23(2), 92–96.
<https://doi.org/10.3766/jaaa.23.2.3>
- Zokoll, M. A., Buschermöhle, M., Abdulhaq, N., Saleh, S., Said, N., Abdulhadi, K., Sellami, F., Hochmuth, S., & Kollmeier, B. (2025). Applicability and Normative Data for an Arabic Matrix Sentence Test for Speech Recognition in Noise. *Cureus*.
<https://doi.org/10.7759/cureus.77062>

APPENDIX

Databases searched include PubMed, ScienceDirect, ResearchGate, and Frontiers. Search terms combined keywords related to speech-in-noise testing (e.g., “speech-in-noise”, “SIN test”, “speech perception in noise”), Arabic language and dialects (e.g., “Arabic dialect”, “colloquial Arabic”, “Lebanese Arabic”), and audiological assessment and validation (e.g., “test development”, “validation”, “normative data”). Studies published between 2000 and 2025 and one in the 1990s were considered, and publications in English and Arabic were eligible.

Table 1.

Representative Arabic SIN studies and key methodological characteristics

Author (Year)	Dialect / Language Used	Sample Size (N)	Age Range	Key Findings
Essawy et al. (2019)	Egyptian Arabic	150 NH + 30 SNHL	Adults (18–60)	Arabic HINT-type sentences reliably distinguished NH and SNHL listeners; adaptive SIN testing sensitive to noise direction; dialect limits generalization.
Elrifaeey et al. (2021)	Egyptian Arabic	60 adults	Adults (18–50)	Arabic QuickSIN provided fast SNR-loss estimation; clinically practical but limited norms and dialect specificity.
Sultan et al. (2020)	Egyptian Arabic	40 adults	Adults (20–55)	Arabic HINT more sensitive and ecologically valid than QuickSIN; both effective but dialect-specific.
Abdel Rahman (2018)	Egyptian Arabic	200 children	6–12 years	Developed Arabic Words-in-Noise test with age norms; strong paediatric tool but highly dialect-dependent.
Kamal et al. (2023)	MSA / Egyptian-based	40 HA users + 20 NH	Adults	Arabic HINT sensitive to microphone directionality and hearing aid benefit; dialect mismatch not controlled.
Asiri et al. (2025)	MSA-based Matrix Arabic	60 (NH + HA + CI)	Adults	Arabic Matrix Test showed good–excellent test–retest reliability; high standardization but low conversational naturalness.
Zokoll et al. (2025)	MSA-based Matrix Arabic	120 NH	Adults	Provided normative data for Arabic Matrix sentences; strong psychometrics but not dialect-representative.
Alokaily et al. (2022)	Saudi / Gulf Arabic	45 adults	Adults	Developed Word-in-Noise screening tool; useful for screening but regionally limited.
Masalski & Morawski (2025)	Arabic Digits (DIN)	>100 (online users)	Adults	Arabic DIN suitable for remote screening; efficient but limited ecological validity for sentence understanding.
Rawas (2023)	Multiple dialects (Review)	N/A	N/A	Highlighted systematic lack of dialect-representative Arabic SIN materials and presence of dialect bias.

*NH= Normal Hearing; HA= Hearing- Aid, CI= Cochlear Implant