University of Southampton

Faculty of Engineering and Physical Science

Institute of Sound and Vibration Research

Assessment of Sensorineural Hearing Loss in Children in the Kingdom of Saudi Arabia

by

Rania Ahmed Alkahtani

Thesis for the degree of Doctor of Philosophy

January 2020

University of Southampton

Abstract

Faculty of Engineering and Physical Science

Institute of Sound and Vibration Research

Thesis for the degree of Doctor of Philosophy

Assessment of Sensorineural Hearing Loss in Children in the Kingdom of Saudi Arabia

by

Rania Ahmed Alkahtani

Sensorineural hearing loss (SNHL) in children could have serious long-term effects if not identified early after onset. To identify SNHL in children early, childhood hearing screening programmes at birth and later ages have been recommended. In the Kingdom of Saudi Arabia (KSA), until recently, there was no nationwide commitment to the universal neonatal hearing screening programme (UNHS). Even if full coverage was achieved, UNHS might be inadequate to identify a high percentage of cases of SNHL, so more options for hearing screening later in childhood are needed.

To better understand the situation of children with SNHL in the KSA, the first two studies of this research estimated the age of identification (AOI) of SNHL in children in the KSA prior to the UNHS, which has not been investigated before, and investigated the characteristics of the affected children. The two cross-sectional studies were either a review of children's medical records (n=1226) or surveys of parents (n=174). The main findings included: (1) a high AOI of SNHL in children (around 3 years old, range from around 0.1-10 years); (2) a strong association between consanguinity, which is known to cause late-onset SNHL, and SNHL was found (in >70% of the children with SNHL); (3) parental concern about child's hearing identified for the first time as a predictor of SNHL in Saudi children; and (4) parents finding it difficult to access audiology clinics. These findings indicate that late-onset SNHL is expected to be prevalent among children in the KSA, and the difficulty of accessing audiology clinics may play a role in delaying the identification of those children. This motivated the development of a hearing screening tool that is suitable to young children, is easily accessible, can be used by non-audiologists, and is sensitive to SNHL.

The developed test, called the Paediatric Arabic Auditory Speech Test (PAAST), which is a speech-in-noise (SIN) test, was inspired by the McCormick Toy Discrimination Test, which suits children from the age of 2 years onwards. It was implemented in a downloadable iPad application that ran the test automatically to widen the possibilities of implementing a hearing screening test. The development of the PAAST included the conducting of five studies to determine the following: (1) pre-recorded speech material equalised for intelligibility; (2) the test-retest reliability of the PAAST with normal-hearing Arabic-speaking adults (n=30); (3) the normal-range and test-retest reliability of the PAAST with normal-hearing Arabic-speaking children (n=40, 3-12 years old); (4) typical results in children with mild to severe SNHL (n=16, 6-14 years old) in the KSA; and (5) the usability and feasibility of the tablet application at home and school by parents (n=26) and teachers (n=24) in the

KSA. The studies also sought to explore the normal developmental trajectory of speech intelligibility and the supra-threshold effects of SNHL in Arabic-speaking children.

The PAAST showed good test-retest reliability when tested with adults, older children (>6-12 years), and young children (3-6 years) (e.g. intra-class correlation coefficient= 0.7, 0.8, 0.7 respectively). It could differentiate between normal-hearing children and hearing-impaired children. A high system usability score (>80/100) was found for parents and teachers. It seems feasible to use the PAAST as a hearing test in schools in the KSA. It was also found that there was probably a developmental age-effect on the performance of Arabic-speaking normal-hearing children on SIN tests and that there were substantial difficulties in performing SIN tests for Arabic-speaking children who have SNHL, which has not been documented previously. In general, it could be concluded that the PAAST provides a useable platform for speech intelligibility testing in noise, and the SIN test seems to provide a useful assessment of speech intelligibility in Arabic-speaking children with SNHL.

Table of Contents

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name: Rania Ahmed Alkahtani

Title of thesis: Assessment of Sensorineural Hearing Loss in Children in the Kingdom of Saudi Arabia

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

Alkahtani R, Rowan D, Kattan N, Alwan N. (2019) Age of Identification of Sensorineural Hearing Loss and Characteristics of Affected Children: Findings from Two Cross-sectional Studies in Saudi Arabia. Journal article published in the International Journal of Pediatric Otorhinolaryngology. https://doi.org/10.1016/j.ijporl.2019.03.019

Alkahtani R, Rowan D, Alwan N. Age of Identification of Sensorineural Hearing Loss in Children in Saudi Arabia. Oral presentation at the 4th Conference of the Advanced Arab Academy of Audiovestibulogy; 2017, 24 – 26th November; Amman, Jordan.

Alkahtani R, Rowan D, Alwan N, Shehabi A. Development of an Arabic Speech-in-noise Test for Children. Poster presentation at the Basic Auditory Science Meeting; 2017, 4 – 5th September; Nottingham, United Kingdom.

Alkahtani R, Rowan D, Kattan N, Alwan N. Age of Identification of Sensorineural Hearing Loss and Characteristics of Affected Children: Findings from Two preliminary studies. Poster presentation at the 13th KFMC Annual Research Symposium; 2018, 20 – 21st November; Riyadh, KSA.

Alkahtani R, Rowan D, Alwan N, Shehabi A. The Children's Audiology Picture Test: An Arabic Speech-in-noise test and iPad Application for Children. Oral presentation at the 9th Emirates Otorhinolaryngology, Audiology and Communication Disorders Congress; 2019, 16 – 18th January; Dubai, United Arab of Emirates.

Alkahtani R, Rowan D, Semeraro H, Alwan N, Shehabi A. The downloadable tablet-based Paediatric Arabic Auditory Speech Test (PAAST). Poster presentation at the Internet and Audiology 4th International Meeting; 2019, 17 – 18th June; Southampton, United Kingdom.

Alkahtani R, Rowan D, Semeraro H, Alwan N, Shehabi A. The Development of the Paediatric Arabic Auditory Speech Test (PAAST). Oral and poster presentation at the Institute of Sound and Vibration Research Human Science Group Away Day; 2019, 24th June; Southampton, United Kingdom. I have been awarded second prize for the best overall contribution.

Alkahtani R. The Development of the Paediatric Arabic Auditory Speech Test for iPad. Magazine article published at the ENT and Audiology News; 2019, May/June issue.

Signature: .... Date: 30-01-2020

Acknowledgements

First, I would like to thank God for giving me the strength, wisdom and patience to complete this PhD thesis. My sincere gratitude goes to my main supervisor Dr Daniel Rowan for his invaluable guidance, support, motivation and encouragement. I appreciate his guidance and support both academically and technically. I also appreciate his patience, his emotional support and his faith that I could complete my PhD. I am grateful to my co-supervisor Dr Nisreen Alwan for her precious guidance, advice, support, enthusiasm and friendship. I appreciate her support both academically and personally. My thanks also go to my co-supervisor Dr Hannah Semeraro for her support, cooperation and encouragement.

I would like to thank Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia and the Global Challenge Research Fund, United Kingdom for funding this work.

Thank you to Dr Nad Kattan, Mrs Noura Almegil, Mr Ali Aljowair, Mr Khalid Alaqeel and Dr Reem Alqahtani for their help in data collection for the first two studies of this PhD research. Thank you to Mr Adnan Alshehabi for his contribution in Study 3, to Mr Ahmed bin Afif for his help with some of the figures included in this thesis, and to my colleagues, Faten and Matt, for their help and support. Thank you to those individuals, adults and children, who participated in my studies.

I would like to thank my parents, Ahmed and Majdah, for their prayers and unconditional support. My deepest thanks go to my beloved and supportive children, Leen, Talal and Khayyal for their great patience at all times, for believing in me and being always by my side. Thank you to my siblings, Rawia, Rakan, Ammar, Samaher, Salman and Sultan for their support throughout my journey.

I would like to thank my second family in Southampton, Shorouq and Mona, for always being there for me. Thank you to all my friends in Southampton, who turned the hard time of being away from home and family into enjoyable and unforgettable memories.

Thank you all for supporting me and believing in me. This work could not be done without you being by my side.

Abbreviations

ANOVA	Analysis of variance
AOI	Age of identification
CHL	Conductive hearing loss
CI	Confidence intervals
dB	Decibel
GP	General practitioner
HICs	High income countries
HL	Hearing loss
ICC	Intra-class correlation coefficient
IQR	Interquartile range
JCIH	Joint Committee on Infant Hearing
KSA	Kingdom of Saudi Arabia
LTASS	Long-term average speech spectrum
MHL	Mixed hearing loss
MOH	Ministry of health
MTT	McCormick Toy Discrimination Test
NH	Normal hearing
PAAST	Paediatric Arabic Auditory Speech Test
pDev	Probability of deviation
PF	Psychometric function
PTA	Pure tone audiometry

RMS	Root mean square
SD	Standard deviation
SDw	Within-subject standard deviation
SEHS	School entry hearing screening
SIN	Speech in noise
SLM	Sound level meter
SNHL	Sensorineural hearing loss
SNR	Signal-to-noise ratio
SRT	Speech reception threshold
SUS	System usability scale
UHL	Unilateral hearing loss
UK	United Kingdom
UNHS	Universal newborn hearing screening
USA	United States of America

Chapter 1 Introduction

Sensorineural hearing loss (SNHL) is usually permanent, as it results from irreversible damages to the cochlea and/or the auditory nerve (Madell and Flexer, 2008). The effect of SNHL on the hearing abilities of affected individuals goes beyond a reduction in the audibility of sounds, as it causes a distortion to received sounds, particularly complex sounds, such as speech (Moore, 2003). This distortion effect makes speech difficult to understand, even when it is made loud enough for individuals to hear it (e.g. through hearing aids). The difficulty is even increased in noisy situations (Moore, 2003), which is common in real-life environments. SNHL is considered one of the main causes of the reduction in communication abilities of affected individuals because humans use their hearing primarily to communicate with each other through speech. The role of back and forth conversations in children's development is critical as it supports children's speech and language development and helps them build the strong relationships they need for their general wellbeing. For children with SNHL, the benefit from conversations may be limited because conversations can be a challenge for them. Apparently, SNHL has serious negative impacts in children's development because, if it is not identified and managed early after onset, it can delay their language and speech development, which in turn could result in social isolation, behavioural problems and academic underachievement. Hence, it reduces the quality of life of affected children (American speech-Language-Hearing Association, 2016). This PhD project therefore focuses on SNHL rather than other types of childhood hearing loss (HL).

It is well-documented that the negative impacts of SNHL in children could be avoided or at least reduced by early identification. Therefore, universal hearing screening programmes, at birth and at later ages, have been recommended (World Health Organization, 2019a). Countries that have implemented universal neonatal hearing screening (UNHS) programmes have reported a dramatic reduction in the age of identification (AOI) of SNHL (Wood *et al.*, 2015; Centers for Disease Control and Prevention, 2017).

In the Kingdom of Saudi Arabia (KSA), the population of interest in this research, the implementation of the UNHS was limited to children born in a small number of hospitals located in the main cities until 2016 (a few months after this PhD project was started), when the first phase of the UNHS started (Ministry of Health, 2016). To date, there has been no nationwide coverage of the UNHS. Even if full coverage of the UNHS was achieved in the KSA, there would still be a need for further hearing screening to identify children who might be missed by the UNHS, fail to follow up after failing the UNHS, develop late-onset SNHL and/or acquire SNHL at later ages. Late in 2018, after the data collection of this PhD project was completed, the Saudi Ministry of Health launched a school screening program which aims to screen students of year 1 and year 4 for several disorders including HL. The program, which targets students enrolled in state schools, is still in its second phase, where only 50% of the students are targeted. Although the implementation of such a program is considered a good step towards better children hearing screening in the country, a large time gap was noticed between the UNHS and the school screening program (6 years). It is therefore considered in this PhD project to investigate the need for other options for children hearing screening that would help in identifying children with SNHL early after onset.

In addition to the limited practice of childhood hearing screening, no data are available in the public domain about the AOI of SNHL in children in the KSA. The availability of information about the current AOI of SNHL in Saudi children allows for evaluating the effectiveness of the UNHS. To provide a better understanding of the situation of children with SNHL in the KSA, the first two studies of this research investigated the AOI of SNHL in children in the KSA prior to the implementation of the UNHS and provided information about the factors associated with SNHL. The last study that looked at the risk factors of HL in children in the KSA, which was more than 10 years ago (Al-Abduljawad and Zakzouk, 2003), found consanguinity and childhood infectious diseases to be prevalent. The profile might have changed since that time as a result of the development of the health system in the KSA in all areas, including health education (Al-Hanawi, 2017). For instance, there was an expected rise in the level of community awareness of different hereditary diseases that could result from consanguineous marriages, such as SNHL. Having a fresh look may be useful to find what predictors of SNHL currently apply and, in the near future, will apply to children in the KSA. This, in turn, would help in planning appropriate services for them in order to identify SNHL early after onset.

The main findings of the first two studies revealed that consanguinity, which could be related to late-onset SNHL, is still practised in the country. This indicates that a percentage of Saudi children are expected to develop late-onset SNHL. Additionally, it was found that parental concern is one of the predictors of SNHL in children in the KSA and that parents reported difficulty accessing audiology clinics. The findings highlighted two important issues: (1) a practical approach is needed to identify children who may develop late-onset SNHL as early as possible; and (2) parents should have their children's hearing assessed immediately in case of suspicion. This motivated the development of a hearing screening tool, which can be accessed and used easily by non-professionals, such as parents, outside of an audiology service setting, in this PhD project. The availability of such a hearing screening tool is expected to allow for screening the hearing of children immediately in case of suspicion without waiting for a visit to an audiology clinic, which may be difficult to access.

The consideration of developing a *flexible* hearing screening tool, which could be used in several settings (e.g. homes, schools, clinics etc.) without the need for special audiology equipment or previous experience in audiology, motivated the development of an automated hearing screening test that could be run on mobile-based devices, which are expected to be easily accessible to almost everyone.

As mentioned previously, SNHL reduces the ability to understand speech, particularly in the presence of background noise (Moore, 2003). In fact, listeners with SNHL usually complain of the difficulty of understanding speech in noisy situations (Davis, 1989; Kramer *et al.*, 1998). It was therefore decided for the screening tool to be sensitive to the communication challenges faced by children with SNHL. This motivated the development of an Arabic speech-in-noise (SIN) test for children. SIN tests simulate real-life listening environments and quantify the abilities of individuals to understand speech in the presence of background noise by estimating the speech reception threshold (SRT) in noise (the lowest level at which an individual can identify speech stimuli 50% of the time) (Madell, 1998). Hence, it assesses the functionality of individuals' hearing.

The developed test, called the Paediatric Arabic Auditory Speech Test (PAAST), was inspired by the McCormick Toy Discrimination Test (MTT), which is a reliable auditory speech test that is used widely in audiology clinics in the United Kingdom to assess the hearing of children from the age of 2 years onwards (McCormick, 1977). The PAAST was implemented on an iPad platform for ease of accessibility without the need for specialist audiology equipment. This may allow for its usage as a hearing-screening tool in both formal and informal settings. For instance, it could be used informally by parents at home, by teachers at school, or by general practitioners and/or nurses at primary care units to screen children's hearing in case of suspicion. It could also be used formally in children hearing-screening programmes. The availability of such an easily accessible hearing-screening tool would help in identifying children with SNHL early after onset because it could be used immediately, in case of a concern, without waiting for a visit to an audiology clinic. This is expected to be useful in a country like the KSA, which has an enormous area (2.15 million km^2) (General Authority for Statistics, 2015) and where people living in rural areas might need to travel long distances to seek tertiary health services, such as audiology services. Hence, it is expected for the development of the PAAST to be a practical step towards better children hearing-screening services in the country.

Following the implementation of the PAAST on an iPad application, it was of interest to explore the usability of the iPad application when administered by non-audiologists (parents and teachers) to overcome any possible problems that could be raised by them before releasing the test for public use. It was also of interest to explore the feasibility of using the PAAST to screen for SNHL in children at schools in the KSA. It was believed that the findings would allow, if it showed that it was feasible for application, the suggestion of an implementation of a preschool hearing screening programme in the country.

The benefits of developing the PAAST are not limited to screening services, however. It would also add to the diagnostic and management services provided to children with SNHL, particularly with the lack of Arabic SIN tests for children. It is expected that the development of the PAAST would add to knowledge in both clinical and research fields. For research, investigating the performance of Arabic-speaking children in SIN tasks, both normal-hearing children and children with SNHL, and investigating the developmental age-effect on the SRT in noise of Arabic-speaking children would add to knowledge, since these topics have not been investigated in previous research. Although it was expected that the performance of Arabic-speaking children in SIN testing would follow the same pattern of children speaking other languages, this issue has not been investigated previously, and thus there is no evidence-based information about it. In addition, the findings of this research regarding the performance of Arabic-speaking children in SIN tasks could be used as a baseline for Arabic SIN tests to be developed in the future. Hence, it seemed that the development of the PAAST would open new areas for future research.

Clinically, the development of the PAAST is expected to aid the diagnosis and management of Arabic-speaking children who have SNHL. This is because pure tone audiometry (PTA), the gold standard of diagnostic hearing tests which is used routinely in audiology clinics in the KSA, assesses the audibility of pure tones but not of complex sounds such as speech. Therefore, the PTA by itself is not enough to assess the communication challenges faced by children with SNHL. To assess the ability of understanding SIN, the use of SIN tests has been recommended and developed in several languages, such as English, Mandarin, Norwegian, Dutch, French etc. (Summerfield *et al.*, 1994; Gelnett *et al.*, 1995; Vaillancourt *et al.*, 2008; Myhrum *et al.*, 2016; Sheikh Rashid *et al.*, 2017; Yuen *et al.*, 2019). SIN tests are not only useful for diagnostic purposes but are also useful for post-diagnostic services such as assessing the effectiveness of an intervention approach (e.g. hearing aids, cochlear implants, auditory training etc.). Hence, it seemed that the benefits of developing the PAAST went beyond screening for SNHL and aiding the identification of SNHL in children early after onset. In fact, it would also improve the assessment and management services provided to Arabic-speaking children with SNHL.

1.1 Structure of the thesis

This research project consists of three main parts (Figure 1.1). The **first part** consists of two studies **(Studies 1 & 2)**, which focused on providing epidemiological information about SNHL in children in the KSA in terms of the AOI with SNHL and the characteristics of the affected children. The findings of the two studies revealed a need for a hearing screening tool that could be easily used and accessible. The **second part** consists of four studies **(Studies 3, 4, 5 & 6)**, which focused on describing the development and evaluation of the PAAST, which is a SIN test that is implemented in an iPad application. The PAAST is believed to help in identifying children with SNHL early after onset. It would also fulfil the need for an Arabic SIN test for children. The **third part** consists of one study **(Study 7)**, which looked at the usability of the PAAST and the feasibility of using it to screen for SNHL in children at schools in the KSA. Figure 1.1 shows a flowchart of the three parts of this PhD project.

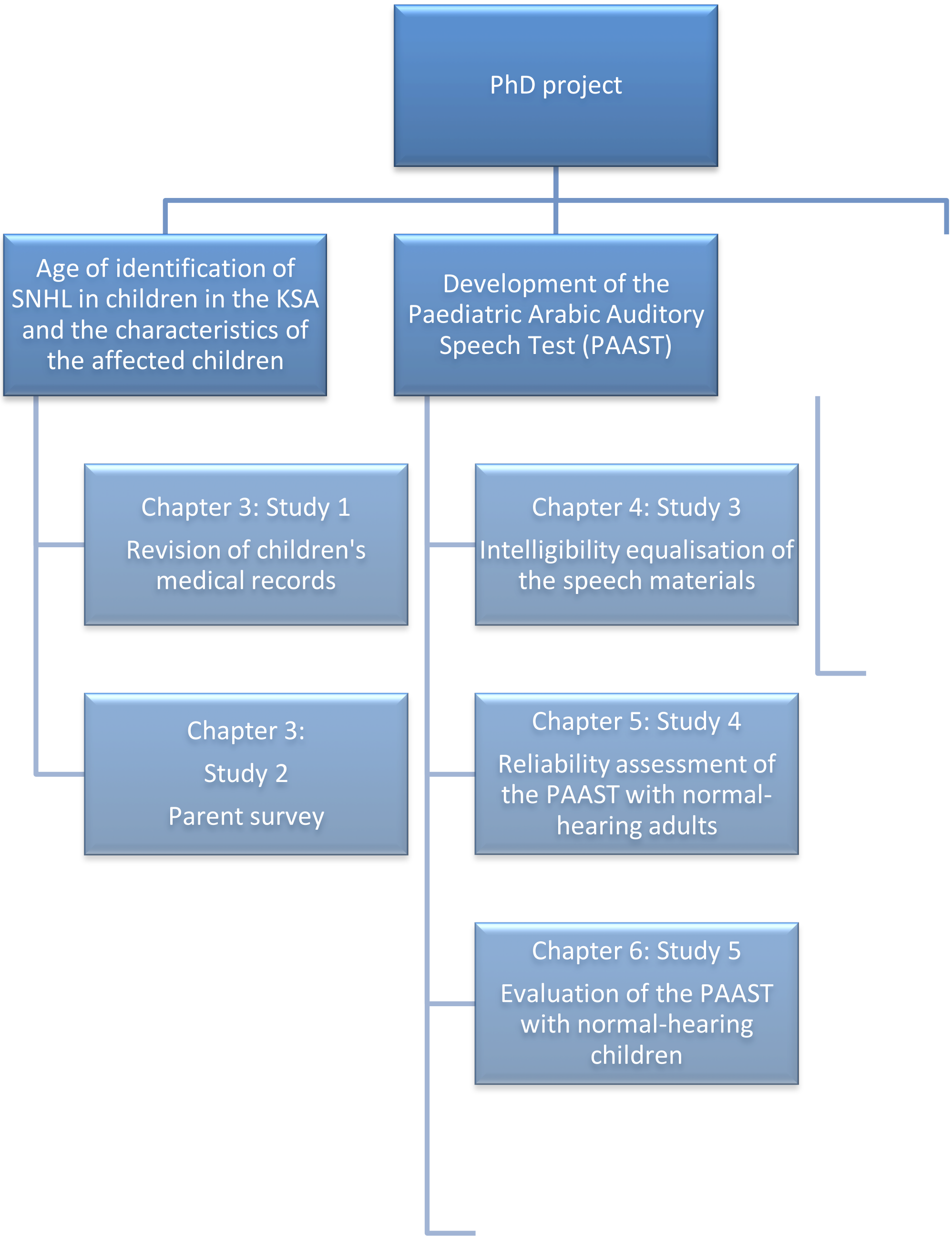


Figure 1.1 Flowchart of the PhD project

The following paragraphs provide an overview of each chapter including the aims of each completed study within each chapter.

Chapter 2 Background and literature review

This chapter consists of three parts. The first part provides a background on the effect of SNHL on hearing abilities, followed by a literature review of the impact of SNHL on children, the role of early identification of SNHL, the practice of infant and childhood hearing screening programmes in high-income countries and an overview of SNHL in children in the KSA. The second part of the chapter outlines the importance of SIN tests, its clinical usage and the elements of developing SIN tests (i.e. speech stimuli, type of noise etc.). The third part of the chapter reviews the use of e-health in audiology, the use of e-health in the KSA, and the implementation of the PAAST in an iPad application. At the end of the chapter, the gaps of knowledge and aims of the research are identified.

Chapter 3 Studies 1 & 2: Age of identification of SNHL in children in the KSA and the characteristics of the affected children

This chapter reports two cross-sectional studies that were carried out with the aims (1) to explore the AOI of SNHL in children in the KSA and (2) to identify the characteristics of the affected children. In Study 1, the medical records of all children who visited audiology clinics in four sites in two cities in the KSA during 2015 were reviewed whereas in Study 2, a questionnaire was completed by carers of children who visited four audiology clinics during a period of three months. Although both studies shared the same aims, Study 1 focused more on the first aim whereas Study 2 focused more on the second aim because of the different methodologies used in the two studies. The work outlined in this chapter was published in the International Journal of Paediatric Otorhinolaryngology (Alkahtani *et al.*, 2019) (https://doi.org/10.1016/j.ijporl.2019.03.019).

Chapter 4 Study 3: Developing, recording and equalising the intelligibility of the speech stimuli of the Paediatric Arabic Auditory Speech Test (PAAST)

This chapter consists of two sections. The first section reports the first stage of developing the PAAST which includes selecting, recording and processing the speech stimuli. The second section reports Study 3, which aimed to measure the intelligibility of the recorded words presented in a stationary speech-shaped noise and equalise the intelligibility by adjusting the amplitude of the words. This stage prepares the speech stimuli to be utilised in an adaptive procedure, which is a method used to efficiently and rapidly estimate the subjects' hearing threshold (Kingdom and Prins, 2010).

Chapter 5 Study 4: Assessing the reliability of the PAAST when utilised in an adaptive procedure (normal-hearing adults)

This chapter reports Study 4, which aimed to explore the test-retest reliability of the PAAST implemented in an iPad application against the test-retest reliability of the PAAST implemented in MATLAB software using tried-and-tested code (considering the MATLAB version as an ideal).

Chapter 6 Studies 5 and 6: Evaluating the PAAST with normal-hearing children and children with SNHL

This chapter reports two studies. Study 5 aimed to explore the test-retest reliability of the PAAST implemented in an iPad application with normal-hearing children and to provide normative data for children. Afterwards, Study 6, which aimed to explore the ability of the PAAST to distinguish between normal-hearing children and children with SNHL was reported.

Chapter 7 Study 7: The usability of the PAAST and the feasibility of using the PAAST in preschool hearing screening in schools in the KSA

This chapter reports Study 7, which looked at the usability of the PAAST implemented in an iPad application when administered by parents and teachers. Additionally, it explored the feasibility of using the PAAST to screen for SNHL in schools in the KSA. Exploring the feasibility and usability of the application was necessary to highlight possible difficulties that may arise when used by individuals who have no experience in audiology.

Chapter 8 General discussion, conclusions and future work

This chapter provides a general discussion about the findings of this PhD project, sums up the main conclusions and proposes possible areas of investigation for future research.

1.2 Contributions to knowledge

The main contributions to knowledge from the current thesis are as follows:

- AOI of SNHL in children in the KSA was provided, which had not been looked at previously. The findings were based on information gathered just before the first phase of UNHS has been established in the KSA. Thus, the reported findings provide a useful database to which the outcomes of the UNHS could be compared. The findings were published in the International Journal of Paediatric Otorhinolaryngology (Alkahtani *et al.*, 2019) (https://doi.org/10.1016/j.ijporl.2019.03.019).

- Parental concern about the child's hearing was identified for the first time as a predictor of SNHL in children in the KSA (Alkahtani *et al.*, 2019). Although it has been identified as a risk factor of SNHL in children in other counties, this was the first study to identify it in the KSA.
- A checklist of risk factors of SNHL in children in the KSA was proposed. It could be used as a first-line screening tool to identify children who are at risk of SNHL for targeted hearing screening purposes.
- An automated Arabic SIN test, the PAAST, was developed and evaluated. Although the development of an auditory speech test is not usually considered a novel contribution to knowledge, the lack of Arabic SIN test adds to the value of the PAAST. The words used in the PAAST are appropriate to Arabic-speaking children in different Arab countries other than the KSA. So, the PAAST can be used to assess the hearing of Arabic-speaking children in different Arab countries. The test was implemented in an iPad application, which eases the accessibility to the test and allows for its usage by non-professionals, such as parents. Moreover, the PAAST serves Arabic-speaking children around the world because the instructions of the test are available in English as well as Arabic. So, English-speaking audiologists can easily use it globally with their Arabic-speaking patients. Additionally, information about the performance of normal-hearing Arabic-speaking children in SIN tests that use adaptive procedure was provided for the first time. Furthermore, the test provides a baseline for further research work that would add to knowledge.
- The work of this research demonstrates that a mobile-device-based hearing test has the potential to be a feasible method of school-based hearing screening run by a non-audiologist. This generates further research proposals such as applying preschool hearing screening in the KSA using the iPad-based hearing test.

Contribution to the research community

- The PAAST could be used widely for clinical and experimental research looking at hearing difficulties children have without being heavily dependent on children's language and cognitive abilities.

1.3 Research activities completed

- **Articles published in peer-reviewed journals**

- Alkahtani R, Rowan D, Kattan N, Alwan N. (2019) Age of Identification of Sensorineural Hearing Loss and Characteristics of Affected Children: Findings from Two Cross-sectional Studies in Saudi Arabia. Journal article published in the International Journal of Pediatric Otorhinolaryngology (https://doi.org/10.1016/j.ijporl.2019.03.019).

- **Presentations at external conferences**

- Alkahtani R, Rowan D, Alwan N. Age of Identification of Sensorineural Hearing Loss in Children in Saudi Arabia. Oral presentation at the 4th Conference of the Advanced Arab Academy of Audiovestibulogy; 2016, 24-26th November; Amman, Jordan.
- Alkahtani R, Rowan D, Alwan N, Shehabi A. Development of an Arabic Speech-in-noise Test for Children. Poster presentation at the Basic Auditory Science Meeting; 2017, 4-5th September; Nottingham, United Kingdom.
- Alkahtani R, Rowan D, Kattan N, Alwan N. Age of Identification of Sensorineural Hearing Loss and Characteristics of Affected Children: Findings from two preliminary studies. Poster presentation at the 13th KFMC Annual Research Symposium; 2018, 20-21st November; Riyadh, KSA.
- Shehabi A, Alkahtani R, Rowan D, Alwan N. The Children's Audiology Picture Test: An Arabic Speech-in-noise Test and iPad Application for Children. Oral presentation at the 9th Emirates Otorhinolaryngology, Audiology and Communication Disorders Congress; 2019, 16-18th January; Dubai, United Arab of Emirates.
- Alkahtani R, Rowan D, Semeraro H, Alwan N, Shehabi A. The downloadable tablet-based Paediatric Arabic Auditory Speech Test (PAAST). Poster presentation at the Internet and Audiology 4th International Meeting; 2019, 17-18th June; Southampton, United Kingdom.

- **Internal presentations**

- Alkahtani R, Rowan D, Semeraro H, Alwan N, Shehabi A. The Development of the Paediatric Arabic Auditory Speech Test (PAAST). Oral and poster presentation at the Institute of Sound and Vibration Research Human Science Group Away Day; 2019, 24th June; Southampton, United Kingdom. I have been awarded second prize for the best overall contribution.

- **Articles published in a professional magazine**

- Alkahtani R. The Development of the Paediatric Arabic Auditory Speech Test for iPad. A magazine article published in the ENT and Audiology News; 2019, May/June issue.

- **Co-supervision of MSc projects**

- Alshehabi, Adnan (2017) Developing a speech in noise test in Arabic, University of Southampton, Faulty of Engineering and Physical Sciences, MSc Dissertation.
- Alarfaj, Hala (2018) Test-retest reliability of the Arabic version of the McCormick Toy Test in a quiet setting, Faulty of Engineering and Physical Sciences, MSc Dissertation.

Chapter 2 Background and literature review

2.1 Chapter overview

The current chapter is divided into three main parts. The first part provides the reader with a background about sensorineural hearing loss (SNHL), which is the type of hearing loss (HL) of interest in this PhD project, in terms of its effect on different psychoacoustic aspects of hearing (Section 2.2.1). Afterwards, a literature review of the impact of SNHL on children's lives, the role of early identification of SNHL in children, the prevalence of SNHL in children in the Kingdom of Saudi Arabia (KSA) and childhood hearing screening practice in the KSA is provided (Sections 2.2.2-2.2.6). The second part of the chapter discusses speech-in-noise (SIN) tests in terms of their importance, clinical usage, and the consideration that should be taken when developing SIN tests (Section 2.3). The third part provides the reader with an overview about e-health, uses of e-health in audiology, uses of e-health in the KSA and the benefits of implementing the Paediatric Arabic Auditory Speech Test (PAAST) in an iPad application (section 2.4). At the end of this chapter, summary and discussion that explains how the three parts of the literature review fit together is outlined and the gaps of knowledge are identified (Section 2.5) followed by an identification of the research aims (Section 2.6) that would address the gaps in the literature and extend the existing research.

2.2 Sensorineural hearing loss in children

The focus in the current PhD project was on the sensorineural type of HL because it is permanent, and it results in serious long-term consequences if not identified and managed early (Madell and Flexer, 2008). Accordingly, the first part of this chapter focuses on SNHL in children. It is acknowledged that most of the evidence in the literature review is not from Arabic-speaking children because of the lack of studies on Arabic-speaking children. However, parts of it are expected to apply to Arabic-speaking children such as the part that discusses the role of early identification of SNHL, which is not strongly language- or culture-dependent. On the other hand, other parts such as the impact of SNHL on children and families, are potentially influenced by cultural issues; thus it might not apply precisely to Arabic-speaking children. In general, it is expected that SNHL would have at least some negative impact on the affected children and families of different language and cultural backgrounds.

2.2.1 Effect of SNHL on hearing abilities

SNHL is caused by pathologies that affect the cochlea and/or the auditory nerve and is usually permanent due to the fact that damages to the cochlea and/or auditory nerve are irreversible. Prior to discussing the effect of SNHL on hearing abilities, it would be useful to provide an overview of how the inner ear (cochlea and auditory nerve) function in response to sounds.

The cochlea, which is a coiled organ filled with fluids, consists of the organ of Corti and the cochlear nerves. The organ of Corti has a basilar membrane covered in tiny hair cells: the outer and inner hair cells. Cochlear nerve fibres are attached to the inner hair cells and meet with one another to form the cochlear nerve, which departs the cochlea and runs up to the auditory cortex in the brain (Martini and Bartholomew, 2003). It is believed that different parts of the basilar membrane vibrate most intensely to a specific frequency of sound, where low frequencies are represented at the apical end and high frequencies are represented at the basal end. Similarly, there is a neural representation to each frequency at a specific place on the basilar membrane where each nerve fibre is tuned to a specific frequency. Thus, the inner ear seems to act as if it contains auditory filters (Fletcher, 1940; Moore, 2003).

When a complex sound such as speech is presented to the ear, each nerve fibre responds to the frequency at which it is tuned. The ear then performs a frequency analysis to break the complex sound into its components in a process called **frequency selectivity**. Damage to the hair cells, as a result of SNHL, causes loss of the sharp tuning of the auditory nerve fibres and widens the auditory filters (Figure 2.1). This widening of auditory filters causes the filters to respond to a wider range of frequencies rather than being specific. This could be perceived by individuals with SNHL as difficulty in distinguishing between sounds of different frequencies even when the stimuli are heard.

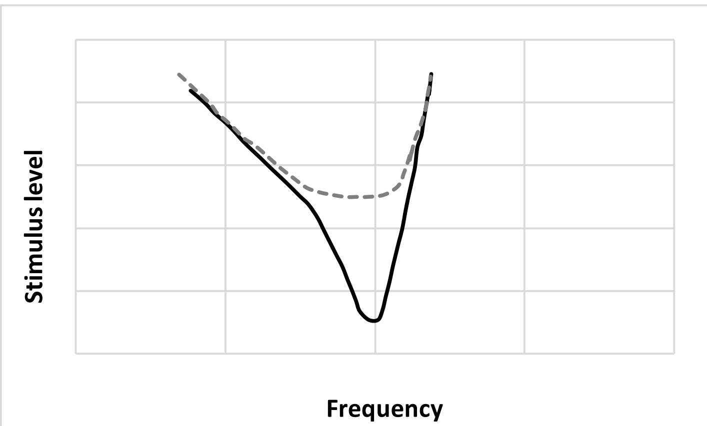


Figure 2.1 An example of the loss of the sharp tuning of auditory nerve fibres. The dashed line represents the widening of the auditory filters

The widening of auditory filters causes difficulty in understanding SIN because in normal ears, when speech is presented in background noise, the filters are assumed to pass the signal and remove a great deal of noise whereas, for individuals with SNHL, the widened filters pass more noise with the same centre frequency as the speech signal, resulting in reducing the signal-to-noise ratio (SNR). Therefore, the affected individuals usually find it difficult to understand SIN as the system does not adequately filter out the unwanted stimuli (masking noise) (Moore, 2003). Several studies looked at the effect of reduced frequency selectivity on the ability to understand SIN in normal-hearing (NH) and hearing-impaired subjects (in the case of hearing impairment, the measurement of the impact was either by simulating widened auditory filters in NH subjects or by testing subjects with SNHL) and showed that the subjects' performance in SIN tasks reduced as the SNR reduced (Dreschler and Plomp, 1980; Florentine *et al.*, 1980; Festen and Plomp, 1983; Horst, 1987; Baer and Moore, 1993, 1994; Xu *et al.*, 2005; Bernstein and Oxenham, 2006; Léger *et al.*, 2012).

Additionally, SNHL results in impairment to the **temporal resolution**, which is the ability to detect changes in stimuli over time, i.e. changes in the temporal envelope. Temporal resolution is essential to understand information carried in complex sounds such as speech because most of the information is carried out in the changing parts of the stimuli rather than in the relatively stable parts. For instance, prosodic information such as intonation, stress and rhythm are usually carried out in the long-term property of the temporal envelop whereas segmental properties of speech, such as consonant articulation and voicing, are carried out in the temporal fine structure (Reed *et al.*, 2009). Temporal resolution is essential for understanding speech, particularly the information contained in the temporal fine structure (Lorenzi *et al.*, 2006; Moore *et al.*, 2006; Hopkins *et al.*, 2008; Hopkins and Moore, 2009; Reed *et al.*, 2009; Moon and Hong, 2014) as it helps in phoneme recognition. Changes in temporal resolution as a result of HL interfere with normal speech perception and phoneme recognition.

Moreover, SNHL results in loss of **cochlear compression**, which allows the cochlea to perceive a wide range of sounds (Moore, 2003). The loss of cochlear compression, which is caused by impaired function of outer hair cells, causes a reduction in the dynamic range and loudness recruitment. The dynamic range is the range between the threshold level and the level where the sounds become uncomfortably loud. Loudness recruitment is an abnormally rapid growth of loudness caused by the reduced dynamic range. Moore and Glasberg (1997) reported that a tone of 100 dB SPL was perceived to be equally loud by both NH individuals and individuals with SNHL. However, the threshold of individuals with SNHL is elevated in comparison to that of NH individuals, which shows the rapid growth of loudness perception in individuals with SNHL. Reduced cochlear compression reduces the speech intelligibility in noise. For instance, in the presence of fluctuating noise, individuals with SNHL do not benefit from the speech information presented in the gaps of noise, called "dip listening" because of the reduced dynamic range. In this case, the most comfortable loudness level of the noise might be close to the hearing threshold, thus the speech information presented in the gaps of noise might be at or below the hearing threshold (Moore, 2003). Therefore, individuals with SNHL would find it more difficult to understand SIN than NH individuals.

As it appears, SNHL impairs different aspects of hearing that play role in analysing and understanding speech. Carhart and Tillman (1970) suggested that SNHL does not only cause a reduction in the threshold of speech in quiet areas but also causes a reduction in the masking efficiency of competing noise when listening in noisy environments. In fact, their suggestion was supported by individuals who have SNHL as they reported that the most difficult thing associated with their HL is the reduced ability to understand speech in the presence of background noise (Davis, 1989; Kramer *et al.*, 1998). In 1978, Plomp developed an SRT model which best described different classes of HL and the effect of each class on understanding speech. According to Plomp (1978), HL of class A is characterised by the reduction in hearing thresholds, which results in a reduction of the ability to hear speech (reduced audibility). However, the reduced audibility itself would negatively affect the ability to understand speech in quiet without understanding SIN because it reduces the audibility of both speech and noise; thus, it does not interfere with the SNR required to understand speech. On the other hand, HL of class D is characterised by distortion to the received stimuli as a result of impaired frequency selectivity, temporal resolution etc. This class of HL decreases the ability to understand SIN because it interferes with the SNR. The effect of SNHL is usually a combination of HL of class A and D as it results in the reduction of audibility and causes distortion to the received message. Hence, even if the intensity of the SNR were increased to increase the audibility of the stimuli, an individual with SNHL would still find it difficult to understand the distorted message.

In sum, SNHL reduces audibility and causes distortion to the received stimuli. Its effect manifests largely in the reduced ability to understand speech in the presence of competing noise. Therefore, considering a measurement of the ability to understand SIN in the assessment of individuals with SNHL is of high importance.

2.2.2 Impact of SNHL on children and their families

It is well recognised that verbal language is normally learned through hearing and listening (American speech-Language-Hearing Association, 2016). Hearing allows external acoustical stimuli to reach the brain through normally functioning ear and auditory nerve and requires no action to be made by the individual. By contrast, the act of listening requires focus and attention (Madell and Flexer, 2008). It is through listening that humans learn. However, in order to listen and learn, humans need to be able to hear normally. Thus, learning crucial skills that are believed to be essential for normal development, such as language, requires normal hearing ability.

The development of the auditory cortex in the brain is not completed until a few years after birth (Paludetti *et al.*, 2012). At this time, the human brain is believed to have high plasticity. In other words, the brain has the ability to build neural networks in response to external stimuli and experiences. Hence, at this point in child's development, exposure to sound and language through hearing would allow the brain to develop normally and build highly organised networks of neuron synapses that support the child in learning life skills, such as language (Paludetti *et al.*, 2012). On the other hand, auditory deprivation caused by HL would reduce the brain's ability to respond to external stimuli. It has been reported that the sensitive period for brain plasticity occurs within the first 5 years of age. Thus, late intervention beyond the sensitive period for children with HL may result in a loss in the brain's ability to recover for the period of auditory deprivation (Paludetti *et al.*, 2012).

SNHL might negatively impact children's communication abilities by delaying the development of language and speech. Reduced communication abilities, in turn, have been widely shown to result in social isolation, behavioural problems and academic underachievement. In addition, SNHL is thought to have an adverse effect on the individual's employment potential (American speech-Language-Hearing Association, 2016).

Although SNHL of any degree is not directly related to behaviour deficits in children, it was reported that SNHL is a high-risk indicator of decreased receptive and expressive language abilities in children, which, in turn, was found to have a direct relation to behavioural difficulties, especially attention difficulties (Stevenson *et al.*, 2010). Thus, SNHL seems to lead indirectly to behavioural problems in children. The following is an example of how the language of children with SNHL differs from the language of children with NH. A child with NH aged 3.5 years could produce around 210 different vocabularies and use sentences of 2.5 words in length when observed playing with his/her parents for 30 minutes. Under the same conditions, a child aged 3 years old with severe SNHL who was taught to use verbal language, could produce only 35 different vocabularies and used sentences that were only 1.5 words in length (Nicholas, 2000).

In addition, SNHL adversely affects children's academic performance by affecting their language development and their ability to understand speech in noisy situations. The United States of America Preventive Services Task Force (USPSTF) stated that the average academic achievement and language levels of primary-school graduates with HL are lower than those of NH fourth-grade children. Additionally, USPSTF reported that the average reading score for graduates with HL is the same as that for fifth-grade NH children (Helfand *et al.*, 2001). It was also found that children who have significant HL >25 dB HL in the better ear perform poorly in elementary school compared to NH children (Järvelin *et al.*, 1997).

SNHL was found to indirectly affect children's ability to learn literacy (reading and writing) by affecting their language development. A child with SNHL would find it difficult to read and write with a poorly developed language and speech system (Dickinson *et al.*, 2006). However, academic success requires more than normal language acquisition. Appropriate communication skills, social interaction abilities and self-esteem levels are all reported to be essential for academic success (Mayer and Wells, 1996). Delayed language, in turn, was found to negatively affect social interaction (Baltaxe, 2001) and self-esteem (Greenberg and Kusché, 1993).

Several studies have investigated the communication and social interaction abilities of children with SNHL in comparison with their NH peers in inclusive educational settings. It was found that children with SNHL seek to initiate communication and social interactions with their peers nearly as often as their hearing counterparts. However, their NH peers were found to ignore them or reject their attempts to communicate, largely due to the limited communication abilities and vocabularies of children with HL (Weisel *et al.*, 2005; DeLuzio and Girolametto, 2011). In a recent cross-sectional study, le Clercq *et al.* (2019) investigated the association between slight/mild SNHL and behavioural problems and school performance in a large sample of children (n= 5355) aged 9-11 years. The findings supported findings from previous studies as they reported that raised hearing thresholds were found to be associated with increased behavioural problems and poorer academic performance. In sum, the extant research works show that SNHL can seriously affect children's normal development.

SNHL affects not only the children who suffer from it, but also their families (Henderson and Hendershott, 1991). To improve their quality of life, family members communicate, have good relationships and enjoy time with one another (Park *et al.*, 2003). Communication difficulties between a child with SNHL and his/her family members were reported to be affected by the hearing abilities of family members, the severity of HL and social support (Hintermair, 2000). More communication difficulties were reported among children with SNHL who have parents with NH. In addition, positive associations were reported between SNHL severity and difficulty of communication (Hintermair, 2000).

It was reported that almost 95% of children who have SNHL are born to NH parents. Thus, it is not unexpected that several reports state that parents of children who were diagnosed with SNHL suffer stressful experiences. The more stress parents experience, the more difficulty children with SNHL face in developing different skills, such as socioemotional and linguistic skills (Hintermair, 2000). Mothers of children with SNHL were also found to be affected by their children's HL. For instance, they were found to touch their children less frequently and were less flexible while playing with them than mothers of NH children. It was also reported that mothers of children with SNHL appeared to need higher levels of social support in order to develop parental behaviour comparable to those of mothers of NH children (Meadow-Orlans and Steinberg, 1993). At the same time, family interaction and support for a child with SNHL was found to be positively related to the child's language development (Moeller, 2000). For this reason, it is essential to support parents in order to help them cope and thus empower their children to cope.

Early identification of SNHL in children was found to help families to release stress and cope better than late identification (Lederberg and Golbach, 2002). Early identification allows family members to seek early support related to coping with the challenges they face while dealing with children with SNHL. This, in turn, facilitates the early development of better relationships between parents and their children with SNHL. On the other hand, late intervention and late family involvement in children's language training were reported to be strongly associated with language delays (Moeller, 2000). In such cases, SNHL may have a negative effect on the family members of children with SNHL, which, in turn, may affect these family members' relations and interactions with their children. In sum, early identification and intervention of SNHL appear to have positive effects on families of children with SNHL.

2.2.3 Role of early identification and intervention of SNHL

Essentially this section outlines the evidence of the benefits of detecting SNHL as early as possible. The proper ways of identifying SNHL early are discussed later in Sections 2.2.4 and 2.2.5. That said, the evidence referred to in this section is not just about detecting SNHL early after onset, it is specifically about the benefits of detecting early-onset SNHL at an early stage. The benefits of detecting later-onset childhood SNHL, after the language critical period, are discussed later.

As discussed in the previous section, SNHL has a negative impact on children's development in many ways. However, early identification (which is the age at which the child is identified with SNHL) and intervention (which is the age at which the child receives intervention such as a hearing aid, cochlear implant etc.) of SNHL in children has shown very promising results. In 1998, a study was conducted to determine the difference in language development between children who were identified with SNHL before the age of 6 months and children who were identified with SNHL after the age of 6 months (Yoshinaga-Itano *et al.*, 1998). Children who were identified earlier were found to have better language development than the other group. Children who were identified later were reported to master approximately 60% of the language abilities of NH children, whereas children who were identified earlier were found to master 80% of the language abilities the NH children had. It is worth mentioning that both groups received early intervention within 2 months of SNHL identification (Yoshinaga-Itano *et al.*, 1998). This indicates that the earlier the age of identification, the better the outcomes of the intervention. It also indicates that even if the intervention took place immediately after the identification, if the age of identification was late, it would still negatively affect the outcomes.

Similar results were obtained by Moeller (2000) who examined the relationship between language development and age at intervention in a sample of 112 children. It was found that children for whom intervention was provided before the age of one year, regardless of the degree of SNHL, showed better language development than children who were provided with intervention after the age of one year. The study also reported that the language development of children with early intervention approximated that of NH children at the age of 5 years. Thus, the study concluded that early intervention results in excellent outcomes, regardless of the degree of the SNHL. In contrast, another study conducted by Wake *et al.* (2005) found that the language abilities of children with SNHL are affected by the severity of HL regardless of the age at confirmation. The study included Australian children aged 7-8 years who were identified as having congenital SNHL. Hearing screening at the age of 8-10 months was available at the time the subjects were born, and intervention took place immediately after identification (Wake *et al.*, 2005). The findings of this study do not agree with those of Moeller (2000), though both studies examined the relations among the age of identification and intervention, the severity of SNHL and language outcomes and included children of similar ages.

The likely explanation for this disagreement in findings might be that, within 6 months of identification, subjects in Moeller (2000) were enrolled in multidisciplinary, family-centred programmes that supported families. Then, the children were enrolled in an extensive intervention programme until the age of 5 years. In contrast, in the study by Wake *et al.* (2005), families were free to choose the mode of intervention provided to their children; thus, interventions varied in both the approach and the intensity of rehabilitation. For this reason, it is difficult to compare the findings of these two studies, since the rehabilitation mode and intensity were not clear in the study by Wake *et al.* (2005). Additionally, the mean age at hearing-aid fitting in Wake's study was 23.2 months, and the mean age at intervention in Moeller's study was 15 months. This delay in intervention might also have affected the outcomes of early identification of SNHL.

Another study investigated the effect of age at intervention on different language skills (Calderon and Naidu, 1999). The language skills of children who were enrolled in interventions before 12 months of age, 13-24 months of age and 25-30 were compared. Children were enrolled in the early childhood home instruction (ECHI) program which is a home-based early intervention program for hearing-impaired children from birth to 3 years old. During the period of the program, an interventionist visited the family at home on a weekly base and used a total communication approach that focused on the language development, auditory skills and speech production of the child. The study found that receptive language, expressive language, speech production and auditory discrimination scores were better among children who received early intervention. Thus, it concluded that the age at intervention is a significant predictor of language skills development. Similar findings were also reported recently by Shojaei *et al.* (2016).

The early identification of SNHL is also related to the development of the social and emotional skills of children with SNHL. It was reported that children in whom SNHL was identified earlier exhibited better social and personal skills than children who were identified with SNHL later (Yoshinaga-Itano *et al.*, 1998). Stika *et al.* (2015) reported that children who were identified early and provided with early intervention including amplification showed developmental outcomes which were similar to NH children. Their results showed greater maternity self-efficacy related to the positive outcomes for those children.

In sum, early identification and intervention of SNHL in children are essential to achieve better outcomes in children's language development, learning abilities and behavioural and social skills.

2.2.4 Universal neonatal hearing screening programme and its limitations

UNHS programmes are applied in many high-income countries (HICs), such as the United Kingdom (UK) and the United States of America (USA) (Uus and Bamford, 2006; Shulman *et al.*, 2010). It focuses on the early identification and intervention of infants with moderate and worse degrees of congenital HL (Public Health England, 2016). Several studies reported that UNHS have a positive impact on speech recognition abilities and language development in children with SNHL (Wolff *et al.*, 2010; Fulcher *et al.*, 2012; Pimperton and Kennedy, 2012; McCreery *et al.*, 2015).

Although it has been demonstrated that UNHS dramatically decreases AOI with SNHL, several factors make the detection of all children with SNHL by the UNHS impossible, even in cases of high sensitivity. First, typical UNHS targets infants with moderate and worse degrees of SNHL. Thus, infants with mild SNHL would pass the screening. Second, it is well recognised that many cases of hereditary SNHL do not manifest at birth. In other words, a child may have normal hearing at birth (and thus pass the UNHS) but may develop SNHL over time. Second, some types of congenital SNHL are progressive in nature. Thus, a child might have slight/mild SNHL at birth (and thus pass the UNHS, since this targets only moderate and worse degrees of SNHL), but then experience deteriorating hearing over time. Third, acquired SNHL, such as SNHL that results from childhood infectious diseases or as a side effect of some medications, does not exist at birth (Madell and Flexer, 2008). Another important point that was highlighted is the high rate (ranging from 3.7% to 65%) of children who have failed the first step of the UNHS programme but also fail to follow up for further investigation (Papacharalampous *et al.*, 2011). Therefore, further hearing screening at later ages is recommended.

It was reported that UNHS identified only 56-59% of school-aged children with SNHL (Nunez-Batalla *et al.*, 2016). Additionally, it was noted that around 0.75/1000 children at preschool age in Shanghai have late-onset SNHL, despite passing the UNHS (Lü *et al.*, 2011). Of these, 0.33, 0.23 and 0.19/1000 were found to have mild bilateral SNHL, moderate bilateral SNHL and unilateral SNHL, respectively. With regard to age, Lü *et al.* (2011) identified SNHL in 1.08/1000, 0.61/1000, 0.56/1000, and 0.76/1000 for children aged 3, 4, 5 and 6 years old, respectively. Only 37.5% of these children exhibited risk factors of SNHL, whereas the other 62.5% did not. Thus, they recommended the establishment of a preschool hearing screening programme in order to support the early identification of those cases with late-onset SNHL that would not be identified otherwise. However, since it was a cross-sectional study, it was not possible to specify whether existing SNHL was late in onset, progressive or acquired.

In another longitudinal study, a cohort of children born in Whipps Cross, London, UK, were followed from birth until their first year at school (Bamford *et al.*, 2007). This study reported that, although UNHS reduced the average AOI with SNHL, there were still children who were not identified until the first year of school (22% with moderate SNHL, 26% with mild SNHL and 18% with unilateral SNHL). The UNHS sensitivity in the study was reported to be 83%, 69% and 46% for moderate and worse, unilateral and mild SNHL, respectively. In general, 3.47/1000 children were identified with SNHL at school entry, with only 1.58/1000 identified by UNHS. In other words, of the children who received UNHS, 1.89/1000 passed the newborn hearing screening but were identified as having SNHL later in life.

Holzinger *et al.* (2016) conducted a cross-sectional study in Carinthia, Austria, in order to estimate the prevalence of significant SNHL among school-age children. The prevalence of significant SNHL in 5 cohorts of school children was estimated to be 2.2/1000. Significant SNHL was defined in the study as a hearing threshold of >40 dB in the better ear. This cut-off point is the same as that considered in the UNHS. However, 50% of children who were identified as having significant SNHL in the study were reported to have passed the UNHS. These children may have developed late-onset SNHL or may have had congenital SNHL but have passed the UNHS regardless. Causes of SNHL in these children were not reported in the study. The mean AOI for those who passed the UNHS or who had not been screened for SNHL at birth was reported to be 4 years old. Based on these results, the authors recommended running ongoing awareness of the importance of early identification of SNHL in children among parents, teachers and physicians. In addition, having hearing screening tests in kindergarten was recommended. This study did not look for mild SNHL, however such inclusion would increase the prevalence of SNHL.

As it appears, UNHS by itself is not sufficient for identifying all children with SNHL, especially when considering unilateral SNHL, mild degrees of SNHL and late-onset SNHL. Further hearing screening at older ages is essential in order to identify these cases as early as possible. The following review highlights the importance of early identification of the cases that would not be detected by the UNHS.

2.2.4.1 Late-onset hereditary SNHL

Hereditary SNHL was reported in the literature to account for 50% of the causes of SNHL, including 30% syndromic and 70% non-syndromic (Bitner-Glindzicz, 2002). As mentioned previously, some hereditary types of SNHL are of late-onset and do not manifest at birth. Several mutations were found to result in late-onset hereditary SNHL that manifests during childhood or even adulthood. For instance, mutations in DFNB2, DFNB8, DFNB10 and DFNB16 were all reported to cause late-onset SNHL (Bitner-Glindzicz, 2002). In addition, it was reported that the GBJ2 mutation, which accounts for 50% of autosomal SNHL, results in late-onset progressive SNHL (Pagarkar *et al.*, 2006). Children affected by late-onset hereditary SNHL would not be identified by UNHS.

In the KSA, one study reported a prevalence of autosomal recessive SNHL of 15.59% (Al-Qahtani *et al.*, 2010), while another study reported a much lower prevalence of 3% (Imtiaz *et al.*, 2011). It is worth mentioning that the study by Al-Qahtani *et al.* (2010) study was conducted in the western region of the KSA, which is home to several migrant families, while the study by Imtiaz *et al.* (2011) study included only Saudi families of Arab tribal origins. Hence, differences in the ethnic backgrounds of the studies' subjects might explain the observed variation in the prevalence of autosomal recessive SNHL. In other words, autosomal recessive SNHL might not be prevalent in Saudi families of Arab tribal origins, but it is still prevalent in Saudi families of other origins.

Other rare mutations were identified in Saudi families and found in the literature. For example, Ramzan *et al.* (2013) identified a mutation in the GIPC3 gene in a Saudi family with 5 hearing-impaired children who were products of a consanguineous marriage. A later study identified an ILDR1 mutation in 3 children of another Saudi family (Ramzan *et al.*, 2014). These children were also products of consanguineous marriage, and they were shown to have bilateral severe to profound SNHL. Ramzan *et al.* (2014) noted that the studied mutations were rare, since it was found in only 1/100 studied families. The W77R mutation was also found specifically in the Saudi and Qatari populations. Moreover, the MYO15A, TMC1, TMPRSS3 and DFNB67 mutations were detected in the KSA and other Middle Eastern countries (Najmabadi and Kahrizi, 2014).

In sum, hereditary SNHL is prevalent in both the KSA and other countries. For this reason, it is important to provide children in these areas with postnatal hearing-screening programmes.

2.2.4.2 Acquired SNHL

Acquired SNHL is well recognised to be prevalent in children. Childhood infectious diseases, known as "TORCH", were reported to cause SNHL in both infants and children. Such diseases include toxoplasmosis, rubella, cytomegalovirus (CMV), herpes simplex virus and others. Meningitis is also a common cause of SNHL. It was reported that 6% of SNHL in children is caused by bacterial meningitis. Other reported causes of acquired SNHL in children include ototoxic drugs, noise exposure and head trauma (Smith *et al.*, 2005).

In a recent study that took place in Mauritania, postnatal causes of SNHL were found in 25.8% of a total of 139 children who were known to have SNHL and were surveyed in order to determine the cause of their HL (Moctar *et al.*, 2016). Infectious diseases were found to be the main causes of acquired SNHL in these children. Meningitis and measles accounted for 12.9% and 2.8% of the causes, respectively. In addition, 3.5% of acquired HL was found to result from the intake of ototoxic drugs.

It was reported that UNHS misses more than half of infants infected with congenital CMV because this virus might be asymptomatic at birth (Fowler *et al.*, 1999). However, 8-15% of infants who do not present with symptoms at birth will develop SNHL at a later age. It was found that 5.2% of children infected with CMV had SNHL at birth, 3.2% developed SNHL by the age of 1 year and another 7% developed SNHL by the age of 6 years. This shows that the majority of cases had late-onset SNHL that would be missed by the UNHS.

In the KSA, 14% of infants and children who had meningitis were found to develop SNHL (Almuneef *et al.*, 1998). In addition, 56% of children aged 12 months to 14 years who were found to be infected with herpes complex were identified as having bilateral SNHL (al Muhaimeed and Zakzouk, 1997). Moreover, in one study, 19% of Saudi people known to have sickle cell disease were diagnosed with SNHL (Al-Dabbous *et al.*, 1996). Although this study included both children and adults (5-40 years old), it reported no difference among age groups in terms of the frequency of SNHL.

Furthermore, in a sample of Saudi children who were examined for toxoplasmosis, 70% of the infected children were found to have bilateral SNHL (al Muhaimeed, 1996). Additionally, in a study conducted in Gizan, a southern city in the KSA, 69% of participants had CMV antibodies at 2 years old, with the highest rate of infection among infants and preschool children (Ashraf *et al.*, 1985). Another study had investigated the aetiology of SNHL in Saudi children and concluded that the most frequent aetiologies were herpes simplex virus type 1, toxoplasma gondii, rubella and CMV (Zakzouk and Hossain, 1994). Thus, childhood infectious diseases seem to be one of the leading causes of SNHL in children in the KSA.

Another well-recognised cause of acquired SNHL in children is the use of chemotherapeutic agents and aminoglycoside antibiotics, which are known to cause ototoxicity. It is known that those drugs might result in permanent high-frequency SNHL. Ototoxic-induced HL could begin immediately after drug administration or exhibit late-onset. It could also be progressive in that deterioration of HL continue to occur over time.

The incidence of ototoxicity was found to be as high as 61% in a sample of 67 children who received chemotherapeutic drugs (Knight *et al.*, 2005). Of the total sample, 17% were referred for hearing aids fitting as a result of their acquired SNHL. In addition, SNHL was found to be progressive in 4.5% of the sample, with progression continuing to occur for 26 months.

Similarly, 67.6% of Japanese children who were treated with cisplatin, a chemotherapeutic drug, were found to develop SNHL. Of these, 17% did not develop SNHL until 2 or more years after taking the last dose of the drug, and another 17% developed progressive SNHL (Yasui *et al.*, 2014).

In the KSA, 7.7% of children who received chemotherapeutic drugs showed SNHL after they received the first dose (Al-Noury, 2011). Another 12% developed SNHL after the second dose of medication. However, children were not followed up further in this study. This is because the aim of the study was to investigate the accuracy of the distortion product otoacoustic emission in screening for SNHL in children who receive cisplatin, a chemotherapeutic drug.

Noise-induced HL (NIHL) is another type of acquired SNHL. It was reported that frequent use of personal stereos is a risk factor for acquired SNHL in school-aged children. A study that investigated the risk factors of SNHL in Australian school-aged children found an association between parents' reports of the use of personal stereos and the presence of SNHL (Cone *et al.*, 2010). Furthermore, in the USA, it was found that 12.5% of school-aged children (6-18 years old) have NIHL and that half of those affected were 7 years old (Niskar *et al.*, 2001). Similarly, findings from a Finnish study reported that 8% of school-aged children (6-7 years old) were found to have NIHL (Haapaniemi, 1995).

In sum, it is clear that acquired SNHL is prevalent in children. Infectious diseases are among the main causes of acquired SNHL in children in the KSA. Additionally, SNHL could be late in onset or even progressive. For this reason, regular hearing assessments for children are highly recommended.

2.2.4.3 Unilateral SNHL

UNHS does not target unilateral sensorineural hearing loss (UHL). It was thought for a long time that UHL has no or at least a minimum effect on children's development (Bess and Tharpe, 1986b). However, several studies have shown that this traditional perception might not be accurate and that UHL might result in more serious problems than expected.

Bess and Thrape (1986a) conducted a study to investigate the effect of UHL on the educational performance of school children. For the purpose of the study, the medical and educational histories of 60 children aged 6-18 years were reviewed. It was found that half of the children with UHL had low academic performance, 35% failed at least one grade in school, and 13.3% required academic assistance. In addition, 20% of children with UHL were noted by their teachers as having behavioural problems. The children were also found to have difficulties in localisation and lower scores in speech recognition than children with normal hearing. Additionally, their speech recognition was found to be affected in both quiet and noisy situations (Bess and Tharpe, 1986a).

Likewise, in a longitudinal study, 46 school children aged 6-12 years who were known to have UHL were assessed annually for 3 years to measure their language abilities, academic performance and behavioural functions (Lieu *et al.*, 2012). It was found that, though the children's language abilities improved over time, no improvement was found in their academic performance. Furthermore, the children's parents and teacher reported problems with socialisation, attention and organisational skills. The authors recommended using individualised educational plans to improve the language skills of children with UHL.

In addition to its effect on school performance, UHL was shown to have an adverse effect on sound localisation ability. It was found that children with UNH have more difficulty localising sounds than their peers of normal hearing (Humes *et al.*, 1980). In addition, HL severity was found to have an effect on localisation ability. In other words, as the degree of UHL increases, the ability to localise sounds decreases. The findings of a study conducted by Bess *et al.* (1986) agreed with the previous findings, showing that children with UHL have more difficulty localising sounds than their peers with NH.

Moreover, it was demonstrated that children with UHL have difficulties understanding SIN. HL severity was found to be related to poor localisation abilities and poor speech recognition abilities in background noise (Ruscetta *et al.*, 2005). Thus, children with UHL might face difficulties at school, where SNR is expected to be low.

The few studies that have investigated the usefulness of UHL intervention have found that the use of FM systems in the classroom increases the speech recognition scores in noise for children with UHL (Kenworthy *et al.*, 1990; Updike, 1994). Additionally, one study found that the use of hearing aids was beneficial for improving the localisation acuity of children with UHL who were identified early, though no benefit was recognised for children identified at late ages, and, in fact, the use of hearing aids was even found to impair localisation acuity in older children (Johnstone *et al.*, 2010). However, this study did not report information on HL severity, and the sample size of 12 children makes generalisation difficult.

Acceptance of hearing aids in children with UHL was investigated. In a sample of 31 children with unilateral SNHL, 67.7% of subjects accepted the use of hearing aids (Georg-August, 2002). Acceptance of hearing aids indicates that the hearing aid benefits the child. Otherwise, the child would refuse its use. However, the same study reported that the acceptance of hearing aids was related to HL severity. As the degree of HL increases, the acceptance of hearing aids decreases. This was explained by the limited benefit of amplification in cases of large gaps in hearing sensitivity between two ears.

UHL was found to be prevalent in the KSA. This could be explained by the high prevalence of childhood infectious diseases within the country. It was reported that 5% of children diagnosed with SNHL have UHL (Al-Abduljawad and Zakzouk, 2003). Bafaqeeh *et al.* (1994) reported a higher prevalence of 7.7% (Bafaqeeh et al., 1994). Moreover, Al-Rowaily *et al.* (2012) reported that 44.4% of all identified children in their study have UHL. Furthermore, Alharbi and Ahmed (2015) reported that 66.6% of SNHL cases identified in their study were unilateral. However, the majority of cases identified by Al-Rowaily *et al.* (2012) and Alharbi and Ahmed (2015) have CHL, which explains the high prevalence of UHL in their findings.

Obviously, it is critically important to screen for UHL in children in the KSA before they start schools. In particular, the extant literature shows that UHL has adverse effects on localisation, understanding speech in noise, and academic performance.

2.2.4.4 Bilateral mild SNHL

The prevalence of mild SNHL among school children shows a variation in the literature. In a cross-sectional study conducted in Malaysia, the prevalence of SNHL in school-aged children was found to be 15%, with mild SNHL accounting for 8.3% (Daud *et al.*, 2010). Furthermore, in the same study, mild SNHL was found to be related to lower academic performance.

Similarly, another study reported that even minimal degrees of SNHL have adverse effects on school children's functional and academic performance (Bess *et al.*, 1998). With regard to minimal SNHL and its relationship to educational performance, this study showed that younger children (grade 3) have lower performance than their matched peers with NH, but that the effect does not translate to older children (grades 6 and 9). However, teachers' responses to the screening instrument for targeting education risk (SIFTER), which assesses school children's academic, communication, attention, behaviour etc., revealed that children with minimal SNHL of all grades had lower scores than their NH peers. In addition, almost a third of children with minimal SNHL (37%) were found to fail a grade during their academic journey. Moreover, these children were found to have behavioural, social, and emotional disturbances, especially in relation to stress, energy and self-esteem.

Children with minimal SNHL were also found to need more listening effort than NH children in order to recognise SIN (Hicks and Tharpe, 2002). Thus, school environments are exhausting for children with minimal SNHL because schools require children to be attentive in considerably noisy situations.

Moreover, several studies have reported language difficulties among children with minimal SNHL. For example, Elfenbein *et al.* (1994) conducted a study designed to assess the oral language skills of children with different degrees of HL. Although some aspects of language, such as semantic and syntactic skills, were found to be similar between children with minimal SNHL and NH children, other aspects, such as pragmatic and articulation skills, were shown to be affected by the SNHL. In addition, children with minimal SNHL reported having difficulties expressing themselves. It is noteworthy that the children themselves suggested that their difficulties were caused by limited vocabulary, speech errors and difficulties in ordering words correctly.

Similarly, in a study in which parents were requested to complete a survey about the language abilities of their children who were known to have minimal SNHL, 40% of parents reported that their children have misarticulations (i.e. the incorrect pronunciation of speech sounds) (Davis *et al.*, 2001). Of those parents, 15% reported that they often did not understand their children's messages.

In the KSA, 33.3% of children with bilateral SNHL were found to have mild SNHL (Al-Abduljawad and Zakzouk, 2003; Alharbi and Ahmed, 2015). An even higher prevalence (57.8%) was reported by Al-Rowaily *et al.* (2012).

Despite these serious consequences of bilateral mild SNHL and UHL, the typical UNHS programmes do not target children with these types and degrees of HL. Additionally, children who would later develop late-onset SNHL of any kind are not detected by these programmes. However, these children deserve early identification and intervention services just as much as their counterparts, who are offered newborn hearing screening testing. Preschool and school-entry hearing screening would help to detect SNHL among these children who would otherwise remain unidentified until a late age.

2.2.5 School entry hearing screening

The practice of school entry hearing screening (SEHS) is applied in HICs, such as the UK and the USA. It has been applied in the UK since the 1930s, although it is now controversial (Bamford *et al.*, 2007). Professionals who have argued against the usefulness of SEHS have reported that primary preventive measures, such as immunisation against infectious diseases, as well as secondary screenings, such as UNHS, have affected SEHS by reducing the incidence of SNHL in preschool children. On the other hand, professionals who have argued for the usefulness of SEHS have reported that even though UNHS has reduced the incidence of SNHL among preschool children, there are still children who might not be identified until the first year of school (22% with moderate SNHL, 26% with mild SNHL and 18% with unilateral SNHL).

Two important studies investigating the practice and effectiveness of SEHS in the UK were found in the literature and are reviewed below.

A. A study conducted by Bamford *et al.* (2007)

In order to determine the effectiveness of SEHS following the implementation of the UNHS in the UK, Bamford and colleagues conducted a questionnaire-based study in 2007. Questionnaires seeking information on the targeted populations, screening conditions, personnel, screening tests, referral criteria, referral rates, yield and coverage of the SEHS programme were distributed to all primary-care trusts, school nursing departments and audiology sites all around the UK. The study achieved a very good response rate (85%), thus making the generalisation of the results reasonable. The intake of the SEHS in state schools and private schools are presented in Table 2.1.

Table 2.1 Intake of SEHS by state and private schools in the UK in percentages

	No SEHS	Targeted SEHS	Universal SEHS
State Schools	6%	7%	87%
Private Schools	51%	29%	20%

Authors of the study failed to explain the reason that some private schools ran only targeted screening or no screening. In addition, the criteria of the targeted children were not clarified. With regard to services for children with known physical, sensory or mental deficits, fewer than half of these services were found to run universal SEHS, nearly a quarter ran targeted SEHS and fewer than a third ran no screenings at all. The authors reported that the reason for the lack of screening was that children who are known to have physical and/or sensory disabilities are expected to have regular medical check-ups, during which their hearing would be tested.

It was reported that, due to the limited use of data-management systems, from which data can be easily retrieved, only a few services offer information on coverage and referral rates. However, from this limited information, 75% of the services were found to have more than 90% coverage, with a median referral rate of 7.9%. Thus, both the percentage of coverage and the referral rate were high. However, it was not clear whether the referred children had CHL, which is common and temporary, or SNHL, which is permanent.

With regard to respondents' view concerning the usefulness of the SEHS, on a scale in which 10 indicated "very useful" and 1 indicated "not at all useful", 70% rated the SEHS with an 8 or higher. Although respondents highlighted the importance of developing a universal protocol in order to conduct more organised services, a very good percentage still found the continuation of the SEHS useful. On the other hand, 28.6% of the services were planning to either stop the screening or change to targeted screening. The study authors did not state the reasons behind such decisions.

Apparently, the intake of SEHS is high in the UK, particularly in state schools, and the majority of schools found it useful and recommend its continuation.

B. A study conducted by Fortnum *et al.* (2016)

Fortnum *et al.* (2016) assessed the effectiveness and cost-effectiveness of SEHS in the UK. Subjects were recruited from Nottingham, where SEHS is applied, and from Cambridge, where no hearing screening at school entry is offered. The authors reported that SEHS was not applied in Cambridge because of both a lack of resources and a lack of protocol and guidance regarding the implementation of the programme. The project consisted of five studies, each focused on a different objective, starting from examining the accuracy of the two screening tests, the pure tone sweeps test and the HearCheck screener, going through investigating the effectiveness of the SEHS and ending with running a cost analysis in order to estimate the cost-effectiveness of the programme.

To assess the effectiveness of the SEHS, retrospective data on all children aged 3-7 years who were referred for audiological assessments through any route except UNHS were collected from the audiology services in Cambridgeshire. The data included the children's age, referral dates, referral sources, types and degrees of HL, types of assessments and interventions, and other information about the staff performing the tests and the equipment used. The collected data were compared to the data that were collected prospectively for age-matched children who were referred for audiological assessment through SEHS in Nottinghamshire.

In Nottingham, where SEHS is applied, the majority of confirmed cases with HL were found to be referred by SEHS (30.8%). Other sources such as ear, nose and throat (ENT) consultants, parents, general practitioners (GPs) and health visitors referred 23.6%, 11.8%, 10.8% and 10.8% of confirmed cases, respectively. On the other hand, in Cambridge, where no SEHS is offered, the majority of confirmed cases were referred by GPs (64.3%). Other resources, such as health visitors and speech therapists, referred 21.4% and 12.2% of cases, respectively. The mean age at the time of referral was 4.7 years for both sites. The mean age of confirmed cases was 4.5 and 5 years in Cambridge and Nottingham respectively. Children with confirmed HL were found to be referred at an older age at the site using SEHS than the site not using SEHS. One possible explanation for this is that, in Cambridge, any suspicion around the age of 4 years results in immediate referral, whereas, in Nottingham, such suspicion would not lead to immediate referral because it is known that the child will undertake a hearing test shortly as part of SEHS. Additionally, the referral rate in Nottingham was found to be lower than that in Cambridge. However, it was reported that Nottingham is more socioeconomically deprived than Cambridge, which would affect both the referral rate and the age at referral. The differences in findings might be caused by differences in the two sites' identification and referral systems. Interestingly, although referring a child who failed SEHS for further investigation resulted in moderate anxiety levels for the parents, 85% of parents of children who were referred strongly agreed with the continuation of SEHS.

In sum of the current practice in the UK, although the application of SEHS is now controversial, uptake rates are still high, and most of the sites recommended the programme's continuation with developing a universal protocol. In addition, the majority of parents are happy with its application. This indicates that general experiences of SEHS in the UK have been successful.

The USA recommends periodic hearing screening tests for preschool and school-aged children (American Academy of Audiology, 2011; American Academy of Pediatrics, 2016; American Speech-Language-Hearing Association, 2019) for the following reasons. First, although the prevalence of congenital SNHL is reported to be 1.4/1000 births, not all children with SNHL are identified by UNHS (American Academy of Audiology, 2011). In addition, not all infants identified by UNHS are being followed up for confirmation. Finally, not all confirmed cases with SNHL are being managed for their HL. Thus, the American Academy of Audiology (2011) recommends applying further hearing screening at older ages. The American Academy of Pediatrics (2016) recommends performing hearing screening tests at the ages of 0, 4, 5, 6, 8 and 10 years old. Additionally, in cases of concern, it is recommended that children of any age be tested for HL.

A study conducted in the USA aimed to investigate the practice of SEHS (Sekhar *et al.*, 2013). They reported that 67% (34/51) of US states apply hearing-screening programmes for children at school age. Of these, 82% require hearing screening, while 18% only recommend it. Of the states requiring hearing screening, 82% apply the hearing screening at different school grades, with 59% applying it even beyond grade 6. The rest of the states (23%) do not apply hearing screening for reasons that were not clarified. Furthermore, different states were reported to use different protocols, and differences in targeted age groups, testing frequencies and intensities were also found.

In sum, the majority of the US states apply school-age hearing screening programmes. These findings suggest that school-age hearing screening is found to be useful in the USA; otherwise, such programmes would likely not be continuously applied, even for older children.

2.2.6 SNHL in children in the KSA

2.2.6.1 Prevalence

In the KSA, it was revealed that the prevalence of HL in children was found to be as high as 13%, with a prevalence of SNHL ranging from 1-4/1000 live births (Al-Abduljawad and Zakzouk, 2003). This is higher than the prevalence of HL among children in HICs which ranges from 1-2/1000 live births (Bamford *et al.*, 2007; Mehra *et al.*, 2009). Despite the seriousness of the problem, only a few studies were found in the literature that investigated HL in children in the KSA. In addition, these studies have yielded widely varying results, with prevalence findings ranging from 1.7% (Al-Rowaily *et al.*, 2012) to 13% (Al-Abduljawad and Zakzouk, 2003).

The followings might explain the variations in findings. First, the study samples have included different age ranges for children with HL. For instance, one study included only kindergarten children aged 4-6 years (Alharbi and Ahmed, 2015), another included school-children aged 6-18 years (Ashoor, 1983) and a third included infants and children aged 6 months to 15 years (Al-Abduljawad and Zakzouk, 2003). Secondly, the studies' sample sizes have differed, with sample sizes ranging from 800 (Ashoor, 1983) to 10,000 subjects (Al-Abduljawad and Zakzouk, 2003). Thirdly, each study was carried out in a different region of the KSA. Thus, the variation in findings might be explained by the geographical and demographical differences in the climate and the socioeconomic status of families living in different areas. Lastly, the time gap between the different studies might also be a factor in the variation in findings. For example, Ashoor's study took place in 1983, while Alharbi and Ahmed's study was conducted in 2015. Improvement in health services over time, which have made family medicine clinics more accessible, is likely to have contributed to differences in findings. Furthermore, the strict implementation of infants' vaccination programmes in the KSA might have decreased the incidence of HL caused by infectious diseases like rubella and measles. For the reasons mentioned above, a comparison of the findings of different Saudi studies in order to estimate the overall prevalence of HL in children in the KSA might not be accurate. Table 2.2 summarises all of the studies found in the literature investigating the prevalence of HL in children in the KSA.

Table 2.2 Summary of different studies investigating the prevalence of HL in children in the KSA

Study	City	Aim	Method	Age of children	Sample size	Prevalence of HL	Prevalence of SNHL
Ashoor (1983)	Dammam	Estimate the prevalence of HL in school children	Survey+ audiological examination	6-18 years	800	7.12%	0.3%
Bafaqeeh *et al.* (1994)	Riyadh	Estimate the prevalence of HL in children	Survey+ audiological examination	0-12 years	6421	7.70%	2.60%
Abolfotouh *et al.* (1995)	Abha	Estimate the prevalence of HL in school children	Audiological examination	6-12 years	974	4.40%	0.30%
Al-Abduljawad and Zakzouk (2003)	4 regions	Estimate the prevalence of SNHL in children	Survey+ audiological examination	6 months - 15 years	9540	13%	1.50%

Study	City	Aim	Method	Age of children	Sample size	Prevalence of HL	Prevalence of SNHL
Al-Rowaily *et al.* (2012)	Riyadh	Estimate the prevalence of HL in school children	Audiological examination	4-8 years	2574	1.75%	0.30%
Alharbi and Ahmed (2015)	Jazan	Estimate the prevalence of HL in kindergarten children	Survey+ audiological examination	4-6 years	1220	3.10%	0.50%

*Prevalence of SNHL out of the total study's sample

In a trial to determine which study offers the most accurate possible estimation, the following comparison between studies was carried out. First, it was noted that Alharbi and Ahmed (2015), who investigated the prevalence of HL in Saudi children aged 4-8 years, excluded children with disabilities from their study sample. Thus, their reported prevalence (1.75%) might underestimate the true occurrence of HL, since it is well recognised that HL could be syndromic and that a high percentage of children who have HL might also have other disabilities (Fortnum and Davis, 1997; Van Naarden *et al.*, 1999; Fortnum *et al.*, 2001). On the other hand, Alharbi and Ahmed (2015) who studied a similar age group (4-6 years old) have included all children and reported a higher prevalence (3.1%). Thus, for this age group, prevalence estimates of 3.1% may be more accurate.

For school-aged Saudi children, the studies by Ashoor (1983) and Abolfotouh *et al.* (1995) also reported quite different prevalence estimates (7.12% and 4.4%, respectively). Nevertheless, both reported estimates are considered high compared to the prevalence of HL in school-aged children in HICs such as the USA, where the prevalence of HL in school children was reported to be 3% (Niskar *et al.*, 1998). On the other hand, this reported prevalence in the KSA is lower than the reported prevalence of HL in school-aged children in low-income countries. For instance, in Karachi and Nigeria, 13.8% and 13.9% respectively, of study samples were shown to have HL (Olusanya, 2001; Hussain *et al.*, 2011). Similarly, in Yemen, 11.6% of primary school children were reported to have HL (Al'shardzhabi and Tsygankova, 2013). Relatively higher prevalence estimates of 20.9% and 20.2% were reported among school children in the Shebin El-Kom district in Egypt and in Iran respectively (Taha *et al.*, 2010; Absalan *et al.*, 2013). It is worth noting that the prevalence estimates reported by Ashoor (1983) and Abolfotouh *et al.* (1995) were based solely on male sample subjects. This might result in an underestimation of the true problem, especially as Al-Rowaily *et al.* (2012) and Alharbi and Ahmed (2015) showed that the majority of cases that were diagnosed with HL were female (71.1% and 56% in the studies of Al-Rowaily *et al.* and Alharbi and Ahmed respectively).

Furthermore, the differences in the observed prevalence between the studies of Ashoor (1983) and Abolfotouh *et al.* (1995) might be explained by the differences in the regions of the KSA in which the studies were conducted as well as the time difference (representing a gap of 12 years). Over time, medical care might have been improved, meaning that subjects in the second study may have had more access to family health services. It is also worth mentioning here that, although these studies were conducted on school children, nothing was reported about the school performance of children identified with HL.

An important study was conducted by Al-Abduljawad and Zakzouk (2003), who used a random sample of Saudi children aged 6 months to 15 years drawn from 4 regions of the KSA (all except the northern region). Another study conducted by Bafaqeeh *et al.* (1994) estimated the prevalence of HL among Saudi children of almost similar age groups (0-12 years); yet, their sample was collected from only one region (Riyadh, central region). Thus, the study of Al-Abduljawad and Zakzouk (2003) represents the only study offering a close estimation of the prevalence of HL among children throughout the KSA. They reported a prevalence of HL among Saudi children of 13%, which is much higher than the prevalence of HL among children in HICs. For instance, in the UK, the prevalence of HL in children younger than 3 years old was reported to be 1.07%, and for children aged 9-16 years old, it was reported to be 2.05% (Fortnum *et al.*, 2001). In addition, a global prevalence of HL among children aged 5-14 years, which was estimated in 29 countries, was reported as 1.4% (Stevens *et al.*, 2013).

Dividing the sample of Al-Abduljawad and Zakzouk (2003) into 4 age groups reveals that most cases of HL were found in the groups of children aged 4-8 years and 8-12 years (36% and 37% of total identified cases respectively). The group of children younger than 4 years old (22%) and the group of children aged 12-15 years (7%) followed these groups. These results are compatible with the results reported by Ashoor (1983), which showed the most identification occurs during children's primary school years. In addition, Al-Abduljawad and Zakzouk (2003) reported that the majority of identified cases in the two age groups (8-12 and 12-15 years old) were female (65.4% & 81.57% respectively). This supports the assumption that the exclusion of females from the studies of Ashoor (1983) and Abolfotouh *et al.* (1995) may have underestimated the prevalence of HL among school children in the KSA. With regard to different regions of the country, most cases were found in the southern and western regions, followed by the eastern and central regions.

Several Saudi studies have recommended hearing screenings for children at preschool age (Abolfotouh *et al.*, 1995; Al-Abduljawad and Zakzouk, 2003; Al-Rowaily *et al.*, 2012; Alharbi and Ahmed, 2015). However, all of the above-mentioned Saudi studies have failed to report whether those children, who were identified with HL in their studies, had been diagnosed with HL prior to their participation in the study or not. Hence, there is no evidence in the literature that the AOI with HL in children in the KSA was investigated in any previous study. However, it is important to provide such information in order to better understand the need for a childhood hearing screening programme.

Although estimating the overall prevalence of HL in children offers a clear picture of the size and seriousness of the problem, it is equally important to determine the prevalence of each type of HL.

This is because some types of HL are more likely to have a permanent effect on children's development if not treated. For instance, CHL is a treatable condition, while SNHL is irreversible (Madell and Flexer, 2008). Thus, SNHL is more serious and requires early detection and intervention.

The prevalence of SNHL among children in the KSA varies among different studies. This prevalence was found to be as high as 2.6% and 1.5% respectively in the works of Bafaqeeh *et al.* (1994) and Al-Abduljawad and Zakzouk (2003). According to Al-Abduljawad and Zakzouk (2003), the prevalence of SNHL among children in the KSA accounts for 1-4/1000 births. This is higher than the prevalence of SNHL among children in HICs, which was reported to be 1-2/1000 births (Bamford *et al.*, 2007; Mehra *et al.*, 2009). However, it is lower than the reported prevalence of confirmed HL in children in low-income countries, which was estimated at 6/1000 births (Störbeck, 2012).

Findings that are comparable to Al-Abduljawad and Zakzouk (2003) were found in Karachi and Kuwait, with a prevalence estimate of 1.1% and 1.6% respectively (Al-Kandari and Alshuaib, 2009; Hussain *et al.*, 2011). However, the sample size of the study conducted in Kuwait was small (159 subjects) (Al-Kandari and Alshuaib, 2009), which made it difficult to generalise the findings. A higher prevalence (3.3%) was reported in Nigeria (Olusanya, 2001). A lower prevalence of SNHL in children in the KSA was reported by Alharbi and Ahmed (2015), who found that 0.5% of kindergarten children were identified with mild SNHL. Furthermore, much lower prevalence estimates were reported by Ashoor (1983), Abolfotouh *et al.* (1995) and Al-Rowaily *et al.* (2012), who found that 0.3% of preschool children have SNHL. The variation in the prevalence of SNHL among Saudi children reported by different studies might be explained by the same reasons discussed previously.

Other studies have investigated HL in newborns in the KSA. This would give an estimation of the prevalence of congenital SNHL in newborns in the KSA. In a study conducted by Maisoun and Zakzouk (2003), 130 neonates who were admitted to a neonatal intensive care unit (NICU) were screened for HL. It was found that 10% of them had moderate to severe SNHL. Another study conducted by Habib and Abdelgaffar (2005) in Jeddah, in the western region, attempted to estimate the incidence of congenital SNHL by screening the hearing abilities of 11,986 not-at-risk babies over a period of 7 years. Their work found an incidence of 0.18% of congenital SNHL among newborn babies. This result is comparable to what was found in Kuwait, where only 2% of healthy newborn babies showed to have SNHL (Al-Kandari and Alshuaib, 2009). However, Al-Kandari and Alshuaib (2007) reported a higher prevalence of HL among high-risk babies (46.67%) than Maisoun and Zakzouk (2003), who identified 10% of at-risk neonates as having SNHL.

Studies from other Middle Eastern countries show high rates of HL among both healthy and at-risk infants. SNHL rate among infants has been found to be 5.34%, 5.12%, 2.80%, 2.21% and 1.37% in Qatar, Kuwait, Nigeria, the Philippines and Jordan, respectively (Olusanya, 2008). These findings are much higher than those in HICs where the reported prevalence of permanent HL in children was 0.1% and 0.11% in UK and USA, respectively (Bamford *et al.*, 2007; Mehra *et al.*, 2009). However, running a comparison of these studies would not produce accurate results because different studies use different criteria for SNHL (i.e. different pass/fail cut-off points and different inclusion/exclusion criteria for bilateral mild and/or unilateral SNHL).

Although SNHL is less prevalent than CHL, the consequences of untreated SNHL are thought to have serious long-term effects on children's development. For that reason, early identification and intervention of SNHL among children in the KSA is crucial. Additionally, identifying the risk factors of SNHL among children in the KSA is of equivalent importance.

2.2.6.2 Risk factors of SNHL in children in the KSA

HL can be either congenital or acquired. In the KSA, the prevalence of congenital HL is reported to be high (Zakzouk and Hossain, 1994). However, acquired HL should not be underestimated, since it also accounts for a significant proportion of identified cases. In a study conducted by Fageeh (2004), in which the aetiology of SNHL was investigated in two schools for deaf children in the Assir region in the southern part of the KSA, 78% of the children were found to have congenital SNHL, whereas 22% had acquired SNHL. Obviously, the prevalence of congenital HL is higher than the prevalence of acquired HL.

Consanguinity, which means related by blood, seems to be a potential risk factor of SNHL. Congenital SNHL was found to be highly related to consanguineous marriages (Shawky *et al.*, 2013). It was stated that almost 8.5% of children or more around the world are products of consanguineous marriages. It is also known that consanguineous marriages are practised largely in Middle Eastern and Arab countries (Shawky *et al.*, 2013) including the KSA due to customs and tribal traditions (Zakzouk *et al.*, 1993). Slightly more than half of the marriages in Arab countries were reported to be consanguineous marriages. In Jordan, Qatar and the United Arab of Emirates (UAE), consanguinity counted for 51.3%, 54% and 50.5% of total marriages, respectively (El Mouzan *et al.*, 2008).

In the KSA, a prevalence of 56% consanguineous marriages was reported, with the highest prevalence (67.2%) occurring in the city of Madinah in the western region, and the lowest prevalence (42.1%) occurring in the city of Al-Baha in the southern region (El-Mouzan *et al.*, 2007).

Other cities, such as Riyadh and Dammam, have 51.3% and 52% consanguineous marriages, respectively (El Mouzan *et al.*, 2008). Rural regions were reported to have a higher prevalence of consanguinity (59.5%) than urban regions (54.7%). In addition, first-cousin relative marriages were found to be more common (33.6%) than other relative marriages (22.4%) (El-Mouzan *et al.*, 2007).

The relationship between consanguineous marriages and hereditary SNHL was investigated in several different studies. In Egypt, for instance, these kinds of marriages were detected in 93.4% of parents of children with HL (Shawky *et al.*, 2013). Similarly, 86.4% of Pakistani children who were diagnosed with HL were found to be products of consanguineous marriages (Sajjad *et al.*, 2008). In Oman, similarly, the prevalence of consanguinity among parents of children with HL was also high (70%) (Al Khabori and Patton, 2008). Likewise, consanguinity was identified in 60.5% of parents of Qatari children identified with HL (Bener *et al.*, 2005).

It was also reported that the incidence of congenital HL in the KSA is notably negatively affected by this kind of marriages. The prevalence of consanguinity and its effect on the prevalence of congenital SNHL was investigated among children in Riyadh, with the result showing that first-cousin consanguinity existed in the parents of 21.1% of affected children and that second-cousin consanguinity existed in the parents of 23% (Zakzouk *et al.*, 1993).

In a longitudinal study by Zakzouk *et al.* (1995), also conducted in Riyadh, 234 subjects aged 7-32 years were followed up for a period of 10 years (1982-1992). Full audiologic examinations including PTA, speech tests, recruitment tests, ABR and CT scan were carried out to assess the hearing sensitivity and to verify the site of lesion. Consanguinity was found in 80.8% of cases with progressive SNHL who were included in the study. Of those consanguineous marriages, 62.8% were first-cousin relatives, and 18% were of second- cousins' marriages. It was stated that symptoms of SNHL were manifested by around 7 years of age and progress with years, indicating late-onset progressive SNHL. In some cases, the progression was rapid whereas in others it was slow. It was also reported that in some families 2-5 members were found to have SNHL. The study has a strength of using a longitudinal design where the subjects were followed up for a long period (10 years). This increases the assertion regarding the accuracy of the findings, which indicate that late-onset progressive SNHL is common among children in the KSA and that it could be related to the practice of consanguinity.

Additionally, Al-Abduljawad and Zakzouk (2003) reported that hereditary factors are the cause of 36.6% of children identified with SNHL. The same study also reported that consanguinity exists in the parents of 47% of children with SNHL, with 19% of these marriages occurring between first-cousin relatives and 28% occurring between second cousins.

In general, the prevalence of hereditary SNHL was reported to be high in the KSA. One study found hereditary factors to be the cause of as much as 66.1% of HL in a sample of Saudi children (Zakzouk and al-Muhaimeed, 1996). Others have suggested that such factors cause more than half of SNHL in children (Bafaqeeh *et al.*, 1994; Smith *et al.*, 2005). The large-scale practice of consanguinity in the KSA was suggested to be the cause of this high prevalence (Bafaqeeh *et al.*, 1994). It was also reported that non-syndromic congenital HL is 2-3 times more in Middle Eastern countries than in the USA and Europe due to the common practice of consanguinity (Najmabadi and Kahrizi, 2014). The literature review above shows that consanguineous marriages play an important role in increasing the prevalence of hereditary SNHL among children in the KSA.

Other risk factors of SNHL were also identified among children in the KSA. For instance, infectious diseases seem to be one of the main factors contributing to the high prevalence of SNHL among children in the KSA. In addition, low socioeconomic status, maternal employment (Bafaqeeh *et al.*, 1994), large family size, low parental education and poor standard hygiene (Ashoor, 1983) were reported in the literature as posing a high risk of SNHL in children in the KSA.

In conclusion, the literature review of the current status of HL among children in the KSA reveals that the prevalence of SNHL among children is high. Hereditary SNHL is common and is highly affected by the practice of consanguineous marriages. Likewise, acquired causes of SNHL are common primarily because of childhood infectious diseases.

2.2.6.3 The practice of childhood hearing screening in the KSA

In the KSA, despite the reported high prevalence of childhood SNHL (Al-Abduljawad and Zakzouk, 2003; Al-Rowaily *et al.*, 2012; Alharbi and Ahmed, 2015), the implementation of the UNHS was limited to children born in a small number of hospitals located in the main cities until 2016, when the first phase of the UNHS started (Ministry of Health, 2016). However, until now, there is no nationwide coverage of the UNHS. Additionally, no data are available in the public domain about the AOI of SNHL in children in the KSA. The availability of information about the current AOI of SNHL in Saudi children would allow for evaluating the success of the recently launched UNHS in reducing the AOI. Additionally, it would provide a picture about the possibility of a need for further hearing screening programmes for children. For that reason, the AOI of SNHL in children in the KSA is considered a topic worth investigation in this PhD project.

Moreover, the last study that looked at the characteristics of children with SNHL was more than 10 years ago (Al-Abduljawad and Zakzouk, 2003), where consanguinity and childhood infectious diseases were found to be prevalent. The profile might have changed since that time as a result of the evolution of the health system in the KSA in all areas, including health education (Al-Hanawi, 2017). For instance, it is expected to see a rise in the level of community awareness of different hereditary diseases that could result from consanguineous marriages such as SNHL. Having a fresh look is useful to find what predictors of SNHL apply to children in the KSA currently and in the near future. This, in turn, would help in planning appropriate services for them. Therefore, this PhD projects looked at the risk factors of SNHL in children in the KSA.

At the time of the literature review of this PhD project was completed, no hearing screening programme for children other than the UNHS was applied in the KSA. However, in 2018, the Saudi Ministry of Health (MOH) launched a school screening programme that targets 50% of year 1 (6 years old) and year 4 students (Ministry of Health, 2019d). The programme screens the students for some disorders, including HL. For screening purposes, nurses visit the school and perform the tests during school time. In 2019, it was announced that 11% of the screened children were identified with HL (Ministry of Health, 2019b). No information was available about the percentage of the identified children in each year (year 1 or year 4), the type and degree of HL in the identified children, whether the HL is bilateral or unilateral, and whether those children were at risk of HL or not. The availability of such information is essential because it is expected for children at these age groups to develop CHL, which is treatable. So, if the majority of the identified children have CHL, the high percentage of children identified with HL would not cause much concern, unlike if the majority of the children have SNHL.

The implementation of the school screening programme is important and recommended to continue and to expand to involve all students of year 1 and year 4 (not only 50% of them) in both state and private schools. However, it is noticed that the time gap between the UNHS and this school hearing screening is large (6 years). There may be a need for a preschool hearing screening to identify children with SNHL early after onset. However, this is unknown for now. In response to that, the current PhD project considered looking at the need for such a hearing screening by providing data about the AOI of SNHL in children in the KSA and the risk factors of SNHL.

2.2.7 Summary

Section 2.2. provided an overview of the impact of SNHL on different aspects of hearing such as the frequency selectivity, temporal resolution, cochlear compression and loudness recruitment. Reduction in such aspects mainly affects listening to complex sounds, such as speech, and causes not only reduction in the audibility of the stimuli but also distorts it. As a result, children with SNHL usually find it difficult to understand speech even when it is audible to them, especially when it is presented in the presence of background noise. The impact of SNHL on children appears in delayed language and speech development, reduction in social activities and interaction, reduced academic performance and increased behavioural problems. The impact extends to reach the families of the affected children as they find it difficult to cope, especially as the majority of children with SNHL are born to NH parents. Fortunately, early identifications of SNHL showed very positive results, as children who were identified early with SNHL and received proper management as early as possible were found to achieve normal language and speech like their NH peers. Therefore, HICs applied childhood hearing-screening programmes for newborns and older children. However, in the KSA, the average AOI of SNHL in children is unknown and, at the time the literature review was completed, the application of childhood hearing screening programmes was limited. This is despite the high prevalence of SNHL among children in the country and the presence of risks factors of SNHL such as the large-scale practice of consanguinity, which could be related to late-onset SNHL. This PhD project therefore considered investigating the AOI of SNHL in children in the KSA and providing a fresh look at the risk factors of SNHL in children in the country in order to look for a possible need for a periodic childhood hearing screening in the KSA.

Based on the literature review, which revealed that consanguinity and childhood infectious disease are prevalent in the KSA, it was expected that there would be a need for a hearing screening to take place between the UNHS and the school hearing screening. It was therefore considered in this PhD project to develop a flexible hearing screening tool that could be accessed and used easily by audiologists and non-audiologists (such as parents and teachers) in several settings (inside and outside the audiology clinics) to widen out the possibilities to how a screen test could be implemented in order to identify children with SNHL early after onset.

In order for the hearing screening tool to identify SNHL in children early after onset, it should be sensitive to SNHL and suitable for young children. It was thought that SIN tests would be useful for that purpose, particularly that it is known that SNHL reduces the ability to understand SIN. Besides, it is expected for auditory speech tests to be more attractive to young children than other auditory tests such as the pure tone audiometry (PTA), particularly if it is designed in a way that attracts their attention and encourage their engagement (e.g. using technology). Therefore, the current PhD project considered the development of a SIN for children in the KSA. The next part of this chapter discusses SIN tests, the clinical importance of such tests and some considerations that need to be taken into consideration when developing a SIN test.

2.3 Speech-in-noise tests

As discussed in Section 2.2.1, the greatest effect of SNHL is manifested as a difficulty to understand speech in the presence of background noise. Plomp and Duquesnoy (1982) reported that a 3 dB loss of SIN causes more disturbance than a 21 dB loss of speech in quiet. Apparently, understanding SIN is a serious challenge that would reduce the quality of life of children with SNHL since real-life environments are noisy most of the time. It is not uncommon to have a noisy listening environment at home. A child with SNHL might struggle to communicate and interact in such a noisy situation. The situation is similar at school not only in playgrounds, where students talk loudly and sometimes scream while playing, but even in classrooms, where ambient noise and poor room acoustics are seen (Crandell, 1993; Knecht *et al.*, 2002; Boothroyd, 2004; Jamieson *et al.*, 2004; Zannin and Zwirtes, 2009; Crukley *et al.*, 2011; Dongre *et al.*, 2017). Noisy conditions are also found in streets, restaurants, parties etc.

The distortion component of HL that is caused by SNHL, unlike the audibility component, is not easily corrected by hearing aids (Plomp, 1978). Therefore, it is essential to assess speech intelligibility in noise of children with SNHL in order to identify children who suffer, measure the extent of their handicap and plan for appropriate management that suits the needs of an individual child. This would help to improve the quality of life of the affected children. The following sections discuss the importance of assessing speech perception abilities using SIN tests, the clinical usage of SIN tests and some consideration that should be taken into account when developing a SIN test for children.

2.3.1 The importance of assessing speech perception abilities using SIN tests

The primary function of the human audition is to communicate with each other, mainly through the perception of speech. Speech understanding abilities, particularly in noise, could not be measured by pure tone audiometry (PTA), the gold standard audiology test, for the following reasons. First, the auditory stimuli used in PTA are pure tones. Pure tones and speech stimuli differ in terms of the characteristics of each stimulus. For instance, pure tones are frequency-specific sinusoidal waveforms, whereas speech stimuli are complex stimuli consisting of different frequencies that change over time. Thus, auditory processing differs for the two types of stimuli (Kuriki and Murase, 1989; Eulitz *et al.*, 1995; Tiitinen *et al.*, 1999; Lin *et al.*, 2005). Also, the effect of SNHL on understanding speech stimuli is more complicated than its effect on pure tones, as described in Section 2.2.1. Second, PTA assesses the ability of an individual to *detect* tones. Thus, it measures the detection threshold, i.e. the audibility aspect of hearing. However, measurement of speech intelligibility is a measurement of the ability to *understand* speech rather than the ability to detect the presence of speech. As discussed in Section 2.2.1., speech intelligibility is affected by both reduced audibility and distortion of the stimuli (HL Class A and D). Although PTA measures the audibility of a tone (Class A HL), it does not provide a measurement of the psychoacoustic properties of sound stimuli, such as frequency selectivity, temporal resolution and cochlear compression. Thus, no information is provided by PTA about the distortion of the sound (Class D HL). Third, PTA is assessed only in quiet, so no information about the individual's performance in the presence of noise is provided.

To evaluate the functionality of the individual's hearing, SIN tasks have been used. In these tasks, the ability of an individual to discriminate between speech stimuli in the presence of noise is measured. The measurement, which is referred to as speech reception threshold (SRT) in noise, identifies the SNR at which 50% of the speech stimuli are identified correctly (Katz *et al.*, 2015).

The correlation between PTA thresholds and SRT in noise was investigated in several studies. The literature showed that, to some extent, there is positive correlation between the two variables (as the PTA threshold worsens, SRT in noise decreases) (Festen and Plomp, 1983; Glasberg and Moore, 1988; Smoorenburg, 1992; Kramer *et al.*, 1996; Beattie *et al.*, 1997; Divenyi and Haupt, 1997; Hagerman, 2002; Smits *et al.*, 2004; Denys *et al.*, 2018). However, the strength of the correlation between the two variables varies widely between the studies, from low correlation such as in Festen and Plomp (1983) (r=0.3) to high correlation such as in Smits *et al.* (2004) (r=0.8). This variation between the studies could be drawn back to different factors such as the sample size, the degrees of HL of the included subjects etc.

Although a correlation could be found between PTA thresholds and SRT in noise, the variation in the SRT between individuals who have similar PTA configuration cannot be fully explained by the PTA thresholds (e.g. Lyregaard, 1982; Killion and Niquette, 2000) because the PTA does not provide measurement to the psychoacoustic properties of sound. To clarify, individuals who have the same configuration of SNHL in PTA might show different SRT in noise because of the different effect of SNHL on the psychoacoustics properties of sound stimuli for each individual. An individual who has a severe reduction in frequency selectivity, for instance, would show worse SRT than another individual who has better frequency selectivity ability even if they have similar PTA thresholds. This has been demonstrated by the findings of Middelweerd *et al.* (1990) who compared the SRTs of individuals who complain of poor ability to understand SIN and had near normal PTA thresholds to the SRTs of a controlled group of individuals who had normal PTA thresholds and had no complaint of hearing difficulty in noise. The findings showed that although both groups had similar PTA thresholds, SRTs were significantly worse in the group who complained of difficulty hearing in noise. Comparable findings were reported by Killion and Niquette (2000) as some individuals who had moderate SNHL (PTA average: 40-60 dB HL) showed no difficulty in SIN tasks whereas others with similar PTA thresholds had an SNR loss of about 15-20 dB. Thus, it seems that the use of PTA thresholds may not reliably predict the effect of SNHL on the ability to understand SIN.

Non-psychoacoustic factors that are not related to auditory abilities such as cognitive abilities, including attention and working memory, also play a role when listening to SIN. These factors could not be accounted for in PTA, which only uses pure tones. According to the ease of language understanding model (ELU), working memory skills could predict individuals' differences in speech recognition in noise (Rönnberg *et al.*, 2013). Working memory is believed to store and process the received speech stimuli in relation to the knowledge stored in the long-term memory. Therefore, it is believed to play a role in speech understanding (Baddeley and Hitch, 1974; Baddeley, 2000; Lunner, 2003; Repovš and Baddeley, 2006; Lunner and Sundewall-Thorén, 2007). Several studies investigated the effect of working memory on speech understanding in noise. For example, Lunner and Sundewall-Thorén (2007) investigated the relative importance of PTA thresholds versus cognitive abilities in predicting SRT in noise in hearing-aid users. They reported that in complex listening situations, PTA thresholds did not contribute to the prediction of speech performance, whereas cognitive test results explained 40% of the variance in speech performance. Similar findings were reported by Sullivan *et al.* (2015) who investigated the effect of noise on working memory and speech understanding in NH children and McCreery *et al.* (2019), who investigated the prediction of speech recognition in noise from auditory, cognitive and linguistic factors in children with HL. In both studies, children with better working-memory skills performed better in SIN tasks than children with worse working-memory skills.

Familiarity with language and vocabulary is another non-psychoacoustic factor that is reported in the literature to contribute to speech understanding in noise (e.g. Nittrouer *et al.*, 2013; Kaandorp *et al.*, 2016). The language abilities of an individual contribute to the ability to recognise SIN as described by the ELU model because the recognition depends on the amount of vocabulary stored in the long-term memory. The lexical restructuring model described the process of the development of vocabulary recognition in noise (Metsala and Walley, 1998). For young children, the whole word is recognised but for older children, where the amount of vocabulary and the similarity between words increases, the process requires differentiation between the phonological representations of similar words in order to be able to distinguish between them. Thus, children with more vocabulary would perform better in recognising words in noise as their ability to accommodate words of similar phonological representations improved. Even for other types of speech stimuli such as sentences, it was found that individuals use the semantic and syntactic clues to recognise the speech stimuli in noise (McElree *et al.*, 2003). Within this context, McCreery *et al.* (2017) looked at the effect of language on speech recognition in noise in NH children. Three types of speech stimuli were used for the purpose of the study: monosyllabic words, word sequences that were semantically and syntactically anomalous, and sentences that were syntactically correct but semantically anomalous. It was found that children with better receptive language skills showed better speech recognition in noise for word sequences and sentences but not for monosyllabic words. However, this was only related to the use of words that were within the vocabulary of the youngest children in the research sample.

In sum, unlike PTA, speech understanding in noise is related to psychoacoustic factors of sound stimuli and non-psychoacoustic factors that are not related to the auditory abilities. Therefore, PTA is not an ideal test to assess the challenges faced by listeners with SNHL; rather, it is better to use auditory speech tests that use SIN tasks.

2.3.2 Clinical usage of auditory speech tests

Auditory speech tests are used to evaluate the functionality of the individual's auditory system by simulating real-life listening environments and quantifying the ability of an individual to understand speech either in quiet or in the presence of background noise (Martin, 1997). They could be used for diagnostic, rehabilitation and screening purposes. In fact, SIN tests would be a good start in the clinic to provide initial information about the patient's difficulty. It could also be used to show the affected individuals or their families the extent of difficulty they face in understanding speech in challenging listening conditions.

SIN tests are usually included within the auditory test battery to support the diagnostic process. They quantify the SNR loss, which is a measure of how well an individual can understand SIN in comparison to individuals who have no difficulty hearing in noise. For example, NH individuals require SNRs about 0 to +2 dB to understand 50% of the perceived speech whereas an individual with 10 dB SNR loss requires the speech to be 10-12 dB louder than the noise to match the performance of NH individuals (Taylor, 2003). Priede and Coles (1976) reported that speech perception tests can be used for differential diagnosis of SNHL. According to them, it could be used to differentiate between SNHL of neural and cochlear origin since neural cases show poorer speech discrimination than cochlear cases. Within the same context of using auditory speech tests in differential diagnosis, Liberman *et al.* (2016) used a SIN test along with other auditory tests in an attempt to diagnose hidden HL in young adults. They included subjects who have normal PTA thresholds up to 8 kHz and divided them into two groups; low-risk and high-risk based on a self-reported questionnaire about noise exposure. They found that the high-risk group showed poor speech recognition in noise regardless of the normal PTA thresholds. This indicates that SIN testing could be used to diagnose SNHL early, particularly the type that results from exposure to loud noises (noise-induced hearing loss).

SIN tests are also used for rehabilitative purposes. Showing the patients or their families the extent of difficulty they have would aid in the counselling to motivate them to commit to the rehabilitation plan given to them. It could also be used to compare the performance of an individual in different conditions. For instance, comparing the performance with and without the use of hearing aids (aided and unaided) would provide information about the usefulness of the hearing aid to the particular individual and would be useful in convincing an individual about the need for hearing aids (Taylor, 2003). In addition, it could be used to compare the performance with and without lip-reading in order to understand how much an individual relies on lip-reading in understanding speech. This would help in counselling the affected individual or the family about where to locate themselves when talking to the child (preferably to face the child if it appears that s/he depends a lot on lip-reading). Moreover, it could be used to compare the performance when different hearing aids are used or when the same hearing aid with different features is used. This would help audiologists to select the best hearing aid with the features that most benefit the particular individual. Furthermore, the successfulness of an auditory training programme could be assessed and monitored by comparing the performance prior, within and following its completion.

Some speech in quiet tests were used for childre's hearing screening purposes purpose. For instance, the verbal auditory screening test (VAST), which includes 12 spondaic words that are presented to the child at different intensity levels (Hamill, 1988), was used for childhood hearing screening in the USA and the UK for a long time (Bamford *et al.*, 2007; McPherson and Brouillette, 2008). Bamford *et al.* (2007) reported in their systematic review of alternative SEHS tests that the VAST is a viable screening option. It was favoured over PTA because children who gave unreliable results with PTA were found to respond well to the VAST (Hamill, 1988). However, it was reported that the VAST has low sensitivity to mild HL and high-frequency HL. Thus, it was not recommended to be used as a hearing-screening test for children (McPherson and Brouillette, 2008).

Similarly, the McCormick Toy Discrimination Test (MTT) was used to screen the hearing of young children. The MTT includes 7 pairs of monosyllabic words represented in toys. The words are presented to the child at several intensity levels, and the child is required to identify the named object (McCormick, 1977; McCormick, 1988). The MTT, which is performed in quiet, showed 100% sensitivity and 94% specificity in detecting CHL in 3-year-old children (Harries and Williamson, 2000). In another study conducted by Bellman *et al.* (1996), the accuracy of the MTT, which was developed for children for whom English is the second language (E2L TOY TEST), was examined. It was found that for the cut-off point of 25dB HL, the sensitivity and specificity were 87% and 90% respectively whereas, for the cut-off point of 20 dB HL, the sensitivity and specificity of E2L TOY TEST were 78% and 94% respectively. Similar results were found for children for whom the first language is English. Thus, it was recommended to use the E2L Toy Test for children's hearing screening in the UK.

Although there is no evidence in the literature that SIN tests were not used previously routinely for SEHS, some SIN tests were developed and evaluated for the purpose of screening the hearing of school-age children such as the internet-based SIN screening test that was developed in Dutch to screen the hearing of children from the age of 5 years onward (Sheikh Rashid *et al.*, 2017) and the digit in noise test that was developed in Flemish and evaluated on children aged 9-16 years (Denys *et al.*, 2018). The SIN test developed by Sheikh Rashid *et al.* (2017) consists of 8 monosyllabic words that are represented in pictures and whereas the SIN test developed Denys *et al.* (2018) consists of 27 combinations of digits triplet. The speech materials of both tests are presented in a speech-shaped noise and an adaptive procedure is used to estimate the child's SRT corresponding to 50% correct responses. The findings of the two mentioned studies revealed that the two tests showed good test-retest reliability and concluded that it can be used for children's hearing screening.

Apparently, SIN tests are of great importance in clinical practice as they support the diagnosis process and aid the rehabilitation management. Hence, it is essential to provide auditory speech tests in different languages within the audiology tests' battery. In the KSA, the use of children's auditory speech tests is limited to speech-in-quiet tests because there was no available SIN test for children in Arabic. The available Arabic auditory speech tests are provided in Table 2.3. At the time of doing the literature review and determining the aims of this PhD project, no SIN test for Arabic-speaking children was published in scholarly articles. Thus, no information was available about the performance of Arabic-speaking children in SIN tests.

Table 2.3 Available Arabic auditory speech tests for children

Speech test	Test condition	Suitable age range	Test stimuli	Form of Arabic language
Pan-Arabic Speech Test (Alusi *et al.*, 1974)	Speech in quiet	Not specified	150 PB monosyllabic words	Standard Arabic
Saudi Arabic Speech Audiometry for Children (Ashoor and Prochazka, 1985)	Speech in quiet	4+ years old	8 lists of 10 PB monosyllabic and disyllabic words	Standard Arabic
Arabic Speech Pattern Contrast (Kishon-Rabin and Rosenhouse, 2000)	Speech in quiet	Not specified	Two lists of monosyllabic word with 8 subsets - each subset includes 12 contrast pairs	Colloquial Arabic spoken in Israel
Arabic Word-in-noise Test[1] (Rahman, 2018)	Speech in noise	5^{+} years old	Three lists of monosyllabic words – lists consist of 25 words	Egyptian colloquial dialect

The most common clinically used speech-in-quiet test for children in the KSA was developed by Ashoor and Prochazka (1985). It consists of 8 lists of monosyllabic and di-syllabic words. Each list consists of 10 phonetically balanced words. The test materials were recorded by a Saudi female speaker and stored on a tape. The test is used to assess the SRTs of children from the age of 4 years old. Although the test is useful to assess children's ability to understand speech in quiet conditions, it does not provide information about their performance in challenging listening conditions (noisy conditions). In addition, the words used are not suitable for children younger than 4 years old. Besides, the words of the test are in the standard form of the Arabic language. Standard Arabic is

[1] The test was published after the data collection of this PhD project was completed.

united and well understood by Arab people in all Arab countries (Saiegh-Haddad and Joshi, 2014) and is mainly used for formal purposes such as in praying, teaching, formal writing, media etc. However, the use of standard Arabic with young children, particularly in new generations, might not be appropriate because children in Arab countries usually learn colloquial language in their first years, which is an informal language that is used for daily communications, has no specific writing style and has different dialects that are specific to geographical regions and demographical backgrounds. Most children in the KSA are exposed to Standard Arabic only when they first join schools, where the language used is the standard language. Although some TV cartoon shows are in standard Arabic, the main form of language that is used by children of a young age is colloquial Arabic.

The only Arabic SIN test that was published in a scholarly article is the Arabic version of the Word-in-noise test (WIN) (Rahman, 2018). It was published after the data collection of this PhD project was completed, consists of three lists of monosyllabic words; each list consists of 25 words. The words are presented with a carrier phrase in a cafeteria noise. The test measures the percentage of correct responses at fixed SNRs ranging from +24 to 0 dB SNR. It is in the colloquial Egyptian dialect; therefore, it is not expected to be familiar to children of other Arab countries, at least not to Saudi children.

Obviously, there is a lack of SIN tests for Arabic-speaking children. No test is available to suit young children in the KSA. Therefore, this PhD project considered the development of an Arabic SIN test that is automated, suitable for children from the age of 3 years old, uses colloquial Saudi Arabic language and implementable in a tablet-based platform.

2.3.3 Considerations for the development of SIN tests

The use of speech materials to assess auditory recognition performance goes back to 1910 when Campbell used nonsense-syllable diagrams to evaluate telephone circuits. A few years later, electronic audiometers, which either have attenuators to vary the presentation level of speech material (Jones and Knudsen, 1924 cited in Wilson and McArdle, 2005) or incorporated a phonograph with recorded digital speech material (Fletcher, 1929) were developed. Afterwards, the development of the principles for constructing and implementing speech recognition tests was provided by Fletcher and Steinberg (1929). Since then, lots of auditory speech tests were developed in quiet and noise.

Currently, several SIN tests for children are available in different languages for clinical use. Some of these tests measure the speech threshold, whereas others measure the maximum speech identification score at supra-threshold intensities. The former is interested in measuring the lowest level at which the individual can identify speech whereas the latter is interested in measuring (in percentages) the maximum performance of the individual when the speech is presented at a fixed loudness level. Both measurements are essential for diagnostic and management purposes such as in hearing-aid fitting. In this PhD project, developing a SIN test that measures children's SRTs was considered. Some considerations that are required to be taken into account when developing a SIN test are described in the following sections.

2.3.3.1 Methods of measurement

To assess the ability to understand SIN, some tests use adaptive procedures, whereas others use the method of constant stimuli (examples are in Table 2.4). Adaptive procedures work by increasing or decreasing the SNRs of the speech stimulus by a fixed intensity amount depending on the individual's response to a previous stimulus. For a known number of presentations, the stimulus level is decreased following a correct response/s whereas it is increased following each incorrect response (Kingdom and Prins, 2010). Thus, it concentrates the stimulus levels on a range that is expected to yield to the smallest standard deviations in threshold estimates. On the other hand, the method of constant stimuli uses a descending presentation level, where the speech stimuli are presented over fixed SNRs (Kingdom and Prins, 2010) at which the listeners achieve scores that range from 0% to 100%. The 50% correct point is then considered the SRT.

The method of constant stimuli is limited by floor and ceiling effects because different individuals have different hearing sensitivities. To clarify, selecting high SNR that is appropriate to some individuals with SNHL, would cause NH listeners to reach a ceiling of 100%, whereas selecting a low SNR that is appropriate for NH listeners would cause some individuals with SNHL to reach floor/near floor effect (very low scores that could reach 0%). Thus, it is impossible to select a fixed SNR that works for a wide range of individuals (Martin, 1997). Even when different SNRs are used, the floor and ceiling effect might appear, particularly when the measured intelligibility deviates from the 50% correct point. On the other hand, this issue is not seen in the adaptive procedure as the SNR is adjusted adaptively to measure a predetermined performance level (e.g. 50%, 70.7% correct responses). Besides, it increases the efficiency of the test and reduces its duration as it concentrates the presentations around the expected threshold based on the subject responses.

The duration of the test should be considered when testing children for two reasons. First, clinicians prefer to use tests that require short duration to complete because of the workload in clinics. Second, it would be difficult for young children to engage in a task for long durations. Using adaptive procedures helps keep young children attentive throughout the test (Elliott *et al.*, 1979). Based on that, it seems that using adaptive procedures is more practical when testing children. Therefore, the SIN test developed in this PhD project uses an adaptive procedure to measure SRT in children.

Table 2.4 Some available SIN tests for children in different languages

Test	Language	Speech stimuli	Noise	Method of speech threshold measurement
McCormick Toy Discrimination test (MCTT) (Lovett *et al.*, 2013)	English	Monosyllabic words	Pink noise – Dutch babble	Adaptive procedure
Bamford-Kowal-Bench Speech-in-Noise Test (BKB-SIN) (Ng *et al.*, 2011)	English	Sentences	Multi-talker babble	The method of constant stimuli
Hearing in Noise Test-Children (HINT-C) (Nilsson *et al.*, 1996; Vaillancourt *et al.*, 2008; Myhrum *et al.*, 2016; Wong *et al.*, 2019)	Available in different languages such as English, Canadian French, Norwegian, Cantonese	Sentences	Speech-shaped noise	Adaptive procedure
Quick Speech-in-Noise Test (QuickSIN) (Vojnovic and Subotic, 2011)	English	Sentences	Multi-talker babble	The method of constant stimuli

Test	Language	Speech stimuli	Noise	Method of speech threshold measurement
Words-in-Noise test (WIN) (Wilson *et al.*, 2010)	English	Monosyllabic words	Multi-talker babble	The method of constant stimuli
Listening in spatialized noise-sentences (LISN-S) (Cameron and Dillon, 2007)	English	Sentences	Distracter stories	Adaptive procedure
Paediatric Speech Intelligibility test (PSI) (Jerger *et al.*, 1980; Jerger and Jerger, 1984; Zheng *et al.*, 2009)	English, Mandarin	Words and sentences	12 speakers	The method of constant stimuli
Paediatric Audiovisual speech test in noise (PAVT) (Arnold *et al.*, 2010)	English	Monosyllabic words	Multi-talker babble	Adaptive procedure
Speech in noise Test (SPIN) (Denys *et al.*, 2018)	Flemish	Digit triplets	Speech-shaped noise	Adaptive procedure
FreeHear (Moore *et al.*, 2019)	English	Digit triplets	Multi-talker babble	Adaptive procedure

2.3.3.2 Types of speech materials

Different speech tests use different speech materials such as isolated sounds, nonsense-syllables, words or sentences. There is controversy in the literature about what speech material to use for speech recognition tests. However, the purpose of the test and the target population are the best determinants of the speech material used.

Isolated sounds are used in some speech tests such as in the Ling six sound test, which assesses hearing across the speech frequencies using the following sounds: /m/, /i/, /a/, /u/, /ʃ/ and /s/. It is useful when testing infants and very young children who are unable to understand and respond to complicated speech stimuli. It is also useful for children who have very severe degrees of SNHL and/or children who are unable to respond to meaningful speech stimuli due to mental illness (Ling, 1978). On the other hand, it would not be useful with older children since it provides limited information about their listening abilities.

Nonsense-syllables are also used in some tests, such as the nonsense stimuli test (Edgerton and Danhauer, 1979). However, it lacks meaning, which makes it unrelated to the vocabulary of the children. In addition, the utterances are usually short, and the sequence of the sounds might not follow the normal sequencing of sounds in the language; thus this might baffle listeners (Martin, 1997). Therefore, nonsense syllables might be difficult to use with children. To test children, it is preferable to use familiar stimuli in order to increase their engagement in the testing, which in turn would give better results. Otherwise, children might not respond or might select a response by chance.

Some researchers recommended the use of sentences because they provide more realistic listening situations as they represent the speech used daily (e.g. Giolas and Epstein, 1963; Jerger *et al.*, 1968; Cox *et al.*, 1987), whereas others recommended the use of words since words are the fundamental units for speech recognition (e.g. Plomp, 2002). The use of sentences has some advantages as it takes into account non-psychoacoustic factors such as cognitive skills (discussed previously in Section 2.3.1). Besides, the individual could use the syntactical and semantic clues in the sentence to recognise the speech, which is usually the case in real-life listening. This is required in some speech tests such as tests that are developed to assess the fitness of duty for military (e.g. Semeraro *et al.*, 2017). In such tests, the contribution of factors other than the auditory function itself should be considered because military subjects are required to use all the available information and clues to understand and follow commands.

However, the case is different when young children are tested, where the involvement of the working memory, for instance, would make the test more complicated to them, particularly if it is administered with competing noise, as they are required to recall several words. Therefore, in such cases the use of words would be more appropriate as it reduces the effect of linguistic context and working memory and forces the listener to depend on the acoustic cues of the stimulus (e.g. Kaandorp *et al.*, 2016; Kaandorp *et al.*, 2017; Nuesse *et al.*, 2018). Besides, the use of sentences increases the duration of the test, which might not be practical with children.

Similar to words, digits are used in some SIN tests to minimise cognitive demand especially as digits are phonetically simple and frequently used (Smits *et al.*, 2004; Hall, 2006; Ozimek *et al.*, 2009; Watson *et al.*, 2012; Zokoll *et al.*, 2012). However, although digits are simple and familiar, they have the disadvantage that they are easily recognisable by only identifying the vowel, at least in English, because each digit has a different vowel (except for five and nine). In Arabic, the digits differ in the number of syllables (two, three and four syllables), which would also make it easily identifiable by the number of syllables.

In the current research, the interest was to develop a paediatric SIN test that is suitable for young children (from the age of 3 years old) in order to allow for its usage in screening the hearing of young children. Apparently, for this age group, the use of words as the speech stimuli would be better than other types of speech stimuli for the reasons described above (meaningful, fewer linguistic clues and depending less on cognitive functions).

2.3.3.3 Types of background noise

Any sound other than the "target" sound is referred to as noise. Background noise interferes and degrades target sounds in a process called "masking", which could have an effect peripherally (in the cochlea) or centrally (at higher levels of the auditory system) (Durlach, 2006).

The peripheral effect of masking occurs as the frequency of the noise overlaps with the frequency of the target sound, thus both the frequency of the target and the noise pass through the same auditory filters in the cochlea. In this case, the representation of the target sound might be degraded. This type of masking is referred to as energetic masking (Shinn-Cunningham, 2008; Leibold, 2017). Steady-state noise such as steady speech-shape noise is believed to produce primarily energetic masking by interfering physically with the target encoding at the peripheral auditory system, at least for young children (Brungart, 2005).

On the other hand, the central effect of masking occurs when the target and the masking noise looks similar at higher levels of the auditory system, causing uncertainty about the masker; thus, making it difficult to identify the target sound (Durlach *et al.*, 2003). This occurs mostly when the noise and speech are perceptually similar, such as in cases where speech is used both as the target and the masking noise. When speech is used as the masking noise, the process of masking differs from that of energetic masking because the masker has fluctuating spectral and temporal configurations and has small pauses between the words. As a result, listeners could benefit from the gaps in the noise to hear glimpses of the target speech and use their cognitive skills to fill in the gaps (based on their stored vocabulary and experience). Masking that results from unwanted speech is called informational masking if the speech was spoken by a few talkers (Brungart, 2001; Freyman *et al.*, 2004). This is because the listener could understand some of the speech that makes the noise. Speech noise that is made of speech spoken by multi-talkers is called multi-talker babble. In this case, the speech that made the noise becomes less speech-like and thus becomes almost incomprehensible. So, the resulting masking, in the case of multi-talker babble, somewhat resembles energetic masking (Freyman *et al.*, 2004) as it becomes less easily confused with the target speech.

Although it seems that the characteristics of informational masking, particularly the presence of glimpses, would make it easier to identify the target speech, it has been found that the segregation between the target and the masking speech might be difficult because of the interference of the cognitive functions as both sounds have linguistic content (e.g. Zhang *et al.*, 2013). Besides, listeners might find it difficult to draw their attention toward the target speech (Carhart *et al.*, 1969), particularly when the number of talkers is few. Large individual differences are seen with informational masking rather than energetic masking (Kidd *et al.*, 2007). This is not surprising since the performance relies not only on hearing abilities but also on cognitive functions.

Findings in the literature suggest that the performance of children varies for different types of noise. For example, informational masking was found to have more effect on children than energetic masking (mainly speech-shaped noise). Hall *et al.* (2002) assessed the ability of NH children aged 5-10 years old and adults in identifying monosyllabic words represented in two-talker masker (informational masking) and speech-shaped noise. They reported that children performed better in speech-shaped noise than in two-talker masker. Similarly, Leibold and Buss (2013), who investigated the performance of NH children aged 5-13 years old and adults in speech intelligibility tasks in two types of noise; two-talker masker and speech-shaped noise, reported that children's performance was better in speech-shaped noise in comparison to their performance in two-talker masker. Large differences were observed in the performance of children in different age groups for the two-talker masker than for the speech-shaped noise. This indicates that children seem less able to benefit from the glimpses to identify some parts of the target speech. It also has an indication that informational masking has longer-lasting developmental effects than energetic masking.

Comparable findings were reported by Corbin *et al.* (2016) who found that the performance of children with speech-shaped noise improved steadily until around the age of 10 years, when the performance reaches adult-like levels, whereas for the two-talker masker the performance did not reach adult-like levels until around 13-14 years old. This indicates that children's ability to understand speech in informational masking continues to mature even after their ability to understand speech in speech-shaped noise matches adults' performance.

On the other hand, other researchers such as Stuart and colleagues (2005, 2006, 2008) reported that children showed a benefit from the gaps presented in an interrupted noise (temporally modulated noise) as their findings revealed better performance with interrupted noise than continuous noise. Similarly, Koopmans *et al.* (2018) found that children performed better in understanding digits in noise when presented in interrupted noise than when presented in continuous noise. It is worth noting that the interrupted noise used in the studies mentioned above is temporally modulated energetic noise. Thus, it differs from speech noise presented in informational masking, which is both spectrally and temporally modulated and has linguistic content. So, it could be hypothesised that children could benefit from the gaps presented in a temporally modulated energetic noise but found it more difficult to benefit from gaps presented in speech noise that has linguistic content.

The information provided above about the speech perception performance of children in different types of noise were all derived from data collected from children who speak languages other than Arabic. There is a lack in the literature about the performance of Arabic-speaking children in speech intelligibility tasks in noise, particularly in SIN test that uses adaptive procedures. Although it would be expected that the performance of Arabic-speaking children would follow the same pattern of performance of children speaking other languages, there is no research evidence of this. It was therefore considered in this research to explore this topic to provide evidence-based information.

For the development of the paediatric Arabic SIN test, it was decided to initially use the standard type of noise, which is the speech-shaped noise. However, in future studies, other types of noise could be used with the PAAST to compare children's performance with different types of noises. This would then help in determining the best type of noise to be used for clinical practice.

2.3.4 Summary

The inclusion of a SIN test in the audiological test battery for children with SNHL is crucial to ensure that they receive appropriate assessments of their communication abilities, which is mainly affected by their reducing ability to understand SIN. The advantages of having SIN tests within the paediatric audiology test battery is not limited to the diagnostic services but also adds a lot to the rehabilitation services provided to children with SNHL. Additionally, it could be used for screening purposes as it is aimed for in this research.

The lack of Arabic SIN tests for children motivated the development of such a test. It is believed that it would add to the hearing screening and diagnostic services provided to children in the KSA. It is also believed to address some gaps of knowledge such as the lack of information about whether SIN tests can be used to screen for SNHL in children, and about the performance of Arabic-speaking children in SIN tests that uses adaptive procedures.

The aim of developing a flexible hearing screening test that could be used in several settings motivated the implementation of the test in a downloadable iPad application. This is believed to widen the spread of the test and to ease its accessibility. The following section reviews the use of e-health in audiology and the use of e-health in the KSA in order to find whether it is acceptable in the KSA. Additionally, the benefits of implementing the hearing test in an iPad application, which is a form of using e-health, is discussed.

2.4 Use of e-health

E-health is defined by the World Health Organization as “the use of information and communication technologies (ICT) for health” (World Health Organization, 2019c). It is used to provide medicine/health services to individuals remotely (i.e. where patients and healthcare providers are in different places). Telemedicine, telehealth and m-health are different types of e-health. Telemedicine is the terminology used to describe providing “clinical care” to patients remotely whereas telehealth is used to describe providing remote “health services” such as patient support and health education (Field and Grigsby, 2002). M-health is the use of smart devices to deliver health services to individuals who are distant from health caregivers via different applications (apps). E-health is an appropriate solution for some clinical challenges such as shortage of staff and limited budget. It provides several solutions such as real-time interaction between patients and healthcare providers, monitoring data provided digitally by patients, offering self-administered health services etc. It is rapidly growing and is used globally by different specialities for different purposes. The World Health Organization (2019c) reported that in the 21st century, e-health should not be considered as an add-on to health services but a central component in the development of health services.

E-health was found to be effective in reducing the cost and improving health services in different specialities. For example, the use of mobile-based apps showed its effectiveness in reducing stress levels in patients who suffer from stress (Hwang and Jo, 2019), in managing and sustaining weight loss in overweight individuals (Alnuaimi *et al.*, 2019), in changing lifestyles and monitoring glucose levels in diabetic patients (Kebede and Pischke, 2019) etc.

2.4.1 Use of e-health in audiology

The World Health Organization (2019a) reported that 466 million persons are living with permanent disabling HL worldwide. Of those, 35 million are children. This large number of hearing-impaired individuals requiring audiology services encourages the use of e-health in audiology in order to make audiology services more accessible and affordable. Several electronic audiology tools were developed for that purpose. For instance, several apps are available to screen hearing using tones (self-administered PTA) (Bright and Pallawela, 2016). However, the accuracy of tone-based apps for hearing screening is limited since different individuals use different ear/headphones, which are not calibrated. Bright and Pallawela (2016) evaluated several tone-based apps and found wide variability in their accuracy depending on the phone and headphones used and the degree of HL. Other hearing-screening tools are available by telephone, websites and smartphone-based apps to assess individuals' ability to understand SIN for adults (e.g. Smits *et al.*, 2004; Spitzlei *et al.*, 2010; Watson *et al.*, 2012; Paglialonga *et al.*, 2014; Potgieter *et al.*, 2016) and children (e.g. Sheikh Rashid *et al.*, 2017; Ramkumar *et al.*, 2018). The advantage of speech-based apps is that they could be used with different phones and headphones (Smits *et al.*, 2004). In 2019, the World Health Organization adopted a speech-based hearing screening app (hearWHO) using a digit-in-noise test (World Health Organization, 2019b).

The use of e-health in audiology showed effectiveness in improving services provided to hearing-impaired individuals. For instance, Smits *et al.* (2006) reported that thousands of people (>159,000) used the telephone-based digit triplets SIN test within the first two and a half years of its release as a national hearing-screening test. To evaluate the outcome, 881 subjects who performed the telephone-based test were surveyed. Almost 97% of the subjects considered the screening test a good initiative, and 95% found the test easy or had little difficulty performing it. Of the subjects who scored "poor", 57% followed the recommendation to visit their GP or audiology specialist. The findings of the study indicate good intake of the public to self-administered hearing tests. Also, the availability of an easily accessible test increased the identification and management for hearing-impaired subjects.

Similarly, Blamey *et al.* (2015) investigated the effectiveness of an online speech reception test in hearing assessment and hearing aid fitting. They found that the use of the online test reduced the cost and improved the hearing-aid fitting. In addition, Barreira-Nielsen and de Almeida Carneiro (2015) investigated the effectiveness of using teleconsultation with elderly hearing-aid users. They found that the use of teleconsultation raised the satisfaction level of using hearing aids among those individuals. Moreover, Ramkumar *et al.* (2018) evaluated the cost and effectiveness of tele-audiology for follow-up diagnostic hearing assessment in rural India. Children who failed hearing screening were followed up remotely for diagnostic assessment. A qualified audiologist performed auditory brainstem response testing in real-time remotely after a trained person prepared the child for the testing. The study concluded that the use of tele-audiology services reduced the travel cost for children's carers and enhanced the chance of diagnosing positive cases, which could be lost by failing to attend follow-up appointments at local clinics.

Within the same context, De Sousa *et al.* (2018) investigated the intake of a smartphone digit-in-noise hearing test (called hearZA). They reported that in a period of a year and 5 months, 24,072 individuals performed the hearing test, of them, 1.275 were children aged 5-15 years. This indicates a high uptake of the hearing test application. Besides, an overall referral rate of 22.4% was reported, indicating that the test is reaching people in need, thus addressing the need of public health.

In sum, it could be concluded that in the age of technology, adapting audiology services to serve the public's needs through e-health is essential, particularly as it showed proper acceptability and intake.

2.4.2 Use of e-health in the KSA

The huge area of the KSA (2.15 million km^2) (General Authority for Statistics, 2015) and the concentration of the healthcare services in urban areas make it difficult for people living in rural areas to access healthcare services. Alfaqeeh *et al.* (2017) reported that the core barrier faced by people in rural areas in Riyadh province in the KSA is the distance to reach primary-care services. If this is the case with primary-care services, it is expected to see greater difficulty accessing audiology services, which are located in the main hospitals in the main cities in the country. Therefore, it is believed that the implementation of e-health would help to enhance healthcare services in the country particularly as the percentage of internet users in the KSA is high (89% of the total population) (Global Media Insight, 2019).

In 2011, the Saudi MOH launched an e-health strategy to support developing comprehensive patient management in the country (Ministry of Health, 2019c). Since then, lots of smart-device health apps were developed to support people and provide healthcare services in different specialities. For example, the MOH released apps that serve patients such as the "Seha" app, which allows people to contact physicians of all specialists for consultation, and the "Mawid" app that allows patients to book their appointments online. Other apps that target MOH's healthcare providers were released, such as the "Ashanak" app, which allows healthcare providers to follow up activities, register for different activities, find vacant jobs etc., and the "Mawared" app that offers self-service to the MOH's staff such as applying for leave (Ministry of Health, 2019a).

Several apps are available to assess and support people with different health conditions such as diet/nutrition, fitness, health/condition tracker, symptom navigator, medication tracker/reminder/manager and chronic disease management apps. According to a survey conducted by Statista (2016), diet/nutrition apps were the leading apps used by Saudi consumers, followed by fitness, whereas chronic-disease management apps were the least used in the KSA. No documented information was found about the use of online/apps audiology services in the KSA. However, hearing-aid providers in the KSA, such as Starkey, Widex and Phonak, offer online hearing-screening tests and mobile apps that could be used in remote hearing-aid fitting and programming. No data is available about the use of these services; however, it is known to the author from experience in audiology clinics in the KSA that people accept these services and use them when needed.

It appears that people in the KSA accept and are willing to use health apps. In a survey conducted by Accenture (2017), it was found that 84% of consumers reported that the use of technology (health apps and wearables) is important in health management. Besides, 81% agreed that the access of their doctors to their e-health records improves the healthcare they receive. Alradhi *et al.* (2019) investigated the acceptance of tele-dental health education in head and neck cancer patients (n=253) in the KSA. The findings showed that patients were happy to be contacted by telehealth providers through technology devices. In addition to the findings that supported the acceptance of patients to receive health services remotely, the study showed that the intake of technology devices is high even among people with low education levels (the education level of almost half of the subjects, 51%, had secondary education or less). It appeared that the majority of the subjects (98%) in the study use technology, 82.1% use smartphones and 51.8% use laptops. It is recognised that the study was conducted on a specific population (cancer patients), therefore the generalisation of the acceptance of telehealth services might not be appropriate. However, the lack of studies investigating the acceptance of e-health in the KSA adds value to the findings of this study.

Albarrak *et al.* (2019) conducted a survey study to investigate the acceptance and willingness of telemedicine among 391 professionals from four hospitals in Riyadh, KSA. The majority of subjects (90%) expressed their willingness to use telemedicine and agreed to the benefits of telemedicine in saving time and money and improving the quality of health services. On the other hand, the findings showed that around half of the subjects (51%) were not familiar with telemedicine tools in various specialities. However, data for this study were collected in 2016; thus, the situation might have changed since then, particularly as knowledge about technology grows rapidly. Other studies reported some challenges to the implementation of e-health in the KSA such as lack of awareness of telemedicine and its benefits, professional resistance, financial issues, technological and privacy challenges (El-Mahalli *et al.*, 2012; Alsulame *et al.*, 2015; Almuayqil *et al.*, 2016). However, the reported studies were investigating the challenges of implementing telemedicine which differs from m-health as described in Section 2.4.1.

It could be concluded that e-health is acceptable and applied in the KSA. Although few studies looked at the acceptance of patients and professionals of using e-health in the KSA, the findings are promising. Raising the awareness about e-health and the available tools would help to increase its intake in the country.

2.4.3 Implementing the PAAST in an iPad application

Implementing the PAAST in an iPad application would ease the accessibility to the test and spread its usage widely in rural and urban areas of the country without the need for special equipment. The availability of the PAAST in an iPad app would ease screening of the hearing of children in case of suspicion by individuals who have no experience in audiology without the need for referrals to audiology clinics. For instance, parents can use it with their children, teachers can use it with students and GPs can use it easily with patients. It could also be accessed easily by audiologists at any audiology clinic regardless of the financial abilities of the hospital/clinic.

The last decade has seen a dramatic increase in the availability of touch devices, such as iPads, in homes and schools that are readily accessible to even very young children. Children (even those of a young age) are familiar with technology (Beschorner and Hutchison, 2013; Wong, 2015) and are stimulated and motivated strongly by the usage of technology (Roschelle *et al.*, 2000). Therefore, it seems that the implementation of the PAAST in an iPad application with a child-friendly interface would make the test more attractive to children by adding fun and enjoyment to the testing procedure.

In the KSA, several projects that support the incorporation of mobile learning through the usage of mobile devices such as smartphones and iPads have been launched by the Ministry of Education in line with growth and development in different areas in the country (Al-Shehri, 2013). Alzannan (2015) found that preschool children in the KSA were motivated and achieved better when they were taught using iPads as opposed to children who were taught using the traditional method. Turkestani (2015) reported that the use of iPads improves the school readiness of Saudi hearing-impaired preschool children in terms of their linguistic, cognitive and social abilities. Sinnari *et al.* (2018) found that the use of technology represented in a tablet application positively influenced the behaviour and cognitive abilities of Saudi children with attention deficit hyperactivity disorder. These findings indicate the appropriateness of using iPad applications with young children in the KSA. It also indicates that iPads are useful tools that could be used to attract young children and increase their motivation.

The advantages of implementing the PAAST in an iPad application is not limited to the KSA, but the benefit is expected to reach children of other Arab countries, particularly with the lack of Arabic SIN tests. Additionally, it was aimed to provide the test's instructions in English as well as Arabic languages; thus English-speaking audiologists who live anywhere around the world and have Arabic-speaking patients can use it easily with them.

Many English-speaking countries have citizens for whom the first language is not English. Previous studies found that subjects whom English is a second language perform worse in English SIN tests than native subjects (Van Wijngaarden *et al.*, 2002; Tabri *et al.*, 2011; Zokoll *et al.*, 2013). Bellman *et al.* (1996), who highlighted the difficulties faced by English-speaking audiologists in using the English speech tests with children for whom English is a second language, reported that it is difficult to develop speech stimuli in all the different languages spoken in those countries. Therefore, they developed a list of English words that are believed to be achieved early by children for whom English is a second language. However, even with this list of words, the children were unlikely to be familiar with all the test's words. Hence, it seems that providing the PAAST in an accessible iPad application would be beneficial for Arabic-speaking children all around the world.

2.5 Summary, discussion and gaps of knowledge

The literature review discussed above shows that SNHL could have serious negative effects on affected children if not identified early after onset. Not all cases of SNHL are manifested at birth because some types of hereditary SNHL are not congenital in nature but, rather, are late-onset. One of the factors related to late-onset SNHL is consanguinity, which is expected to cause autosomal recessive SNHL because consanguinity is known to have a high chance of pairing recessive alleles of SNHL (Shawky *et al.*, 2013). SNHL could also be acquired late in childhood as a result of environmental factors such as childhood infectious diseases, treatment with ototoxic medication etc. It is therefore recommended to apply childhood hearing screening not only at birth but even at later ages.

The literature review also shows that consanguinity is practised highly in the KSA, the population of interest in this research, and that the majority of the children identified with SNHL are products of consanguineous marriages (Zakzouk *et al.*, 1993; Zakzouk and Hossain, 1994; Zakzouk *et al.*, 1995). Thus, it is expected that at least some of those children developed late-onset SNHL. However, many years have passed since the last study that looked at risk factors of SNHL in children in the KSA, and the health system in the KSA has changed since then. Therefore, it was of interest to take a fresh look at the risk factors of SNHL in children in the KSA to better understand the current situation.

The literature review also shows that no study has investigated the AOI of SNHL in children in the KSA and that the practice of childhood hearing screening was not initiated in the KSA until recently. The coverage of the UNHS is not yet fully achieved, and the newly launched school hearing screening, which is only partially implemented, is administered at the age of 6 years. This indicates that children who develop late-onset SNHL or who might be missed by the UNHS might not be identified until they reach the age of 6 years. By that time, the period of the high plasticity of the brain would have been missed, and children would have experienced a relatively long period of at least partial auditory deprivation (Paludetti *et al.*, 2012), which would result in serious negative impacts on their general wellbeing. Hence, there is a need for a practical solution that could help in identifying children with SNHL in the KSA early after onset. In addition, the lack of information about the AOI of SNHL in children in the KSA highlighted the need for providing such information, which would be useful in the future to assess the effectiveness of the UNHS.

Based on this part of the literature review, three gaps in knowledge were identified: (1) the AOI of SNHL in children in the KSA is unknown; (2) the last investigation of the risk factors of SNHL in children in the KSA was more than 10 years ago, and thus no updated information is available; and

(3) it is unknown whether there is a need for a preschool hearing screening, which may be advisable, especially since there is large time gap between the UNHS and the school-based hearing screening in the KSA.

The second part of the literature review focuses on SIN tests, which measures how individuals function with their remaining hearing, because SNHL is known to cause difficulty in understanding speech particularly in the presence of background noise. The literature review shows a lack of Arabic SIN test. In fact, no published Arabic SIN test that suits young children in the KSA was found in the literature. Besides, there is a lack of information about the performance of Arabic-speaking children in SIN tests, particularly those that use adaptive procedures. It is also unknown whether SIN tests can be used to screen for SNHL in children. This was considered the fourth gap of knowledge identified in the literature review.

Based on these gaps in scholarly knowledge, it was decided to develop an Arabic SIN test that could be used to screen for SNHL in children and would add to the audiology assessment services provided to Arabic-speaking children, as nothing currently exists. The author's motivation to put research into practice and to implement a positive change in the audiology practice in the KSA inspired the idea of implementing the developed SIN test on an iPad platform to ease the test's accessibility without the need for specialist audiology equipment. This led to the third part of the literature review, which looked at the use of e-health in audiology and in the KSA. The literature review revealed that the use of e-health in the KSA is acceptable and applied in the KSA, which encouraged the implementation of the developed SIN test in an iPad application and an exploration of its usability and feasibility to screen for SNHL in children at schools in the KSA.

The **gaps of knowledge** identified in the literature review are thus as follows:

- The AOI of SNHL in children in the KSA is unknown
- The last investigation into the risk factors of SNHL in children in the KSA was more than 10 years ago, and thus no updated information is available
- It is unknown whether there is a need for a hearing screening to take place at a preschool age, which may be desirable, especially since the time gap between the UNHS and the school-based hearing screening in the KSA is large (6 years)
- It is unknown whether SIN tests can be used to screen for SNHL in children. Additionally, there is no Arabic SIN test for children that is suitable for children in the KSA.

These gaps are addressed in this PhD project, as described in the research aims below.

2.6 Research aims

The **three main aims** of this PhD project are as follows:

1. To provide data about the AOI of SNHL in children in the KSA and the characteristics of affected children to create a database for implementing and evaluating national childhood hearing screening programmes
2. To develop an Arabic SIN test and assess it in term of its reliability, its ability to screen for SNHL in children, and its usability when performed by parents and teachers
3. To explore the feasibility of using the developed test to screen for SNHL in Kindergarten children at schools in the KSA.

The three aims of the research were achieved by working on the following **objectives**:

Objectives to achieve Aim-1:

- Identify the AOI of SNHL in children in the KSA (Chapter 3, Studies 1 and 2)
- Identify the characteristics of the affected children (Chapter 3, Studies 1 and 2).

Objectives to achieve Aim-2:

- Select and record the speech materials of the Arabic SIN tests; the Paediatric Arabic Auditory Speech Test (PAAST) (Chapter 4)
- Equalise the intelligibility of the speech materials of the PAAST (Chapter 4, Study 3)
- Explore the performance of NH Arabic-speaking adults in SIN tests using the PAAST (Chapter 5, Study 4)
- Explore the test-retest reliability of the PAAST with NH adults (Chapter 5, Study 4)
- Explore the performance of NH Arabic-speaking children in SIN tests using the PAAST (Chapter 6, Study 5)
- Explore the test-retest reliability of the PAAST with NH children and provide preliminary normative data (Chapter 6, Study 5)
- Explore the ability of the PAAST to distinguish between NH children and children with SNHL (Chapter 6, Study 6).

Objectives to achieve Aim-3:

- Explore the usability of the PAAST when used by parents and teachers (Chapter 7, Study 7)
- Explore the feasibility of using the PAAST to screen kindergarten children's hearing at schools in the KSA (Chapter 7, Study 7).

Chapter 3 Studies 1 and 2: Age of identification of SNHL in children in the KSA and the characteristics of the affected children

3.1 Overview of the chapter

The first aim of this PhD project was to provide data on the age of identification (AOI) of sensorineural hearing loss (SNHL) in children in the Kingdom of Saudi Arabia (KSA) and the characteristics of the affected children. This chapter outlines how the first aim was achieved and provides recommendations that would help to improve the current situation for the benefit of children with SNHL in the KSA.

3.2 Introduction

Either genetic or environmental factors cause childhood SNHL. In countries with well-developed health services, genetic abnormalities, which account for at least 50% of SNHL cases worldwide, are considered the main risk factor for childhood SNHL. However, environmental factors such as childhood infectious diseases, which account for approximately 25% of SNHL cases worldwide, are considered the main risk factors for childhood SNHL in low-income countries where health services are not well developed (Schrijver, 2004; da Silva Costa *et al.*, 2017). In the KSA and other countries in the Middle East, the high practice of consanguinity (Zakzouk *et al.*, 1993; Al Khabori and Patton, 2008; El Mouzan *et al.*, 2008; Sajjad *et al.*, 2008) may result in increasing the cases of genetic SNHL (Bafaqeeh *et al.*, 1994).

Symptoms of genetic SNHL does not necessarily manifest at birth; it could rather develop at later ages. This is seen with children who are products of consanguineous marriages because consanguinity increases the probability of pairing recessive alleles of SNHL (Shawky *et al.*, 2013) and causes autosomal recessive SNHL, which could be of late-onset. As discussed in the literature review (Chapter 2, Section 2.2), the identification of SNHL in children early after onset is crucial to avoid or at least to reduce its negative impact on the children's quality of life. Previous studies of the risk factors for childhood SNHL in the KSA found that consanguinity and infectious diseases are common (Al-Dabbous *et al.*, 1996; al Muhaimeed, 1996; al Muhaimeed and Zakzouk, 1997;

Almuneef *et al.*, 1998; Al-Abduljawad and Zakzouk, 2003). This raises the expectation that late-onset SNHL might be prevalent among children who have SNHL in the KSA.

To identify SNHL early after onset, childhood hearing screening programmes have been applied successfully in high-income countries (HICs) (Wood *et al.*, 2015; Centers for Disease Control and Prevention, 2017). In Middle Eastern countries, the implementation of hearing screening programmes for children is very limited, perhaps due to other economic priorities and low levels of awareness of the seriousness of childhood SNHL (Olusanya *et al.*, 2007). In the KSA, the implementation of the universal neonatal hearing screening programme (UNHS) was limited to children born in a small number of hospitals located in the main cities until 2016, when the first phase of the UNHS started (Ministry of Health, 2016). Until now, there has been no nationwide cover of the UNHS. However, even if full coverage was achieved, UNHS might be inadequate to identify a high percentage of cases of SNHL (reasons were discussed in Chapter 2, Section 2.2.4). Therefore, more options for hearing screening later in childhood are needed.

In 2018, a school screening programme was launched by the Saudi Ministry of Health (MOH) (Ministry of Health, 2019d). The programme screens year 1 and year 4 students for several disorders, including hearing loss (HL). Currently, the programme is in its second phase, where only 50% of students enrolled in state schools are targeted (Ministry of Health, 2019b). The implementation of such a programme is encouraging as it is expected to identify more children with HL. However, because children in the KSA do not enrol in year 1 until the age of 6 years, the identification of SNHL at this age would still be considered late, particularly for children who have early-onset SNHL (at birth) or who develop late-onset SNHL at early childhood (e.g. at 2-3 years old). This is because the sensitive period for brain plasticity, at which the brain has the ability to build neural networks in response to external stimuli and experiences, occurs within the first 5 years of age. Hence, late identification beyond the sensitive period for children with SNHL may result in a loss in the brain's ability to recover for the period of auditory deprivation (Paludetti *et al.*, 2012). It is therefore recommended to apply children's hearing screening within the first 5 years of age and to immediately screen the hearing of the child in case of suspicion in order to identify SNHL early after onset (American Academy of Audiology, 2011; American Academy of Pediatrics, 2016; American Speech-Language-Hearing Association, 2019).

The most recent study that looked at risk factors of SNHL in children in the KSA was more than 10 years ago (Al-Abduljawad and Zakzouk, 2003), and the profile of risk factors may have since changed following amendments in health services and public health awareness, which are improving rapidly in the KSA (Al-Hanawi, 2017). Therefore, it was considered in this research to provide a fresh look at the risk factors of SNHL that applies currently to children in the KSA.

In addition, no data are available in the public domain about AOI of SNHL in children in the KSA. Several studies investigated the prevalence of SNHL in children in the KSA but not the age at which the affected children were identified (Abolfotouh *et al.*, 1995; Al-Abduljawad and Zakzouk, 2003; Al-Rowaily *et al.*, 2012; Alharbi and Ahmed, 2015). The availability of information about the current AOI of SNHL in Saudi children would allow for evaluating the effectiveness of the UNHS, which has been recently initiated in the KSA, in reducing the AOI. Additionally, it would show a possible need for periodic hearing screening programmes for children.

3.2.1 Aims

The aims of Studies 1 and 2 were as follows:

1. To provide initial data on the AOI of SNHL in Saudi children, to inform the evidence base for implementing and evaluating national hearing-screening programmes
2. To take a fresh look at potential predictors of childhood SNHL in Saudi children

Achieving these two aims contributes to achieving the first aim of this PhD research.

3.3 Methods

To achieve the first aim of this PhD project, two studies were conducted. Study 1 used a cross-sectional design that gathered medical record data from all children who visited several audiology clinics in two main cities—Riyadh (the capital city) and Dammam (a major city in the eastern region)—during 2015. Study 2 used a self-report survey of carers of a different group of children with SNHL. Both studies gathered information on the AOI and characteristics of children. While Study 1 provided a larger sample size than Study 2, Study 2 was able to provide more information about the characteristics of children with SNHL due to limitations in what is documented in medical records. Consequently, Study 1 focused more on the AOI, whereas Study 2 focused more on risk factors.

3.3.1 Study 1

In this cross-sectional study, all hospitals/centres with audiology clinics, which are known to the researcher, in different regions of the KSA were contacted (there was no reliable information about the total number and locations of audiology clinics in the KSA). Of the contacted hospitals/centres, 11 were invited to participate. Those 11 hospitals, which are located in different regions of the KSA (central, eastern, and western), were selected because of their use of electronic medical records for their patients. Emails were sent to the head of the audiology unit in each hospital, explaining the research aims and requesting their participation. Reminder emails were sent to non-responders 4 weeks later. Final reminder emails were sent 6 weeks after the first email. Three hospitals did not respond, and 4 hospitals declined participation for reasons related to workload or inability to release patient data for research purposes. Subsequently, 4 hospitals/centres were included in the study; 3 in Riyadh (the capital city) and 1 in Dammam (a major city in the eastern region). These are the Security Forces Hospital, King Fahad Medical City, Gouf Speech and Hearing Centre and John Hopkins Aramco Healthcare.

For the purpose of the study, audiologists in each audiology clinic retrospectively reviewed the electronic medical records of all children aged 0-10 years who attended their clinics during the year 2015 (n=1224). Some of those children were attending for their first audiology appointment and some for follow up. Extracted information for this study included age at the day of data collection, gender, results of hearing tests, degree and type of HL if it existed, AOI of SNHL or mixed hearing loss (MHL), and any other past/present medical condition that may be related to HL. The chance of double counting was very low since three of the included hospitals were governmental hospitals, each of them serving a particular population.

Of the four included hospitals/centres, two have newborn hearing screenings for all newborn babies, which started in 2006 (8.6 years before data collection); one has targeted hearing screening for babies in the Newborn Intensive Care Unit (NICU), which started in 2010 (5 years before data collection); and the fourth has no hearing screening. Information on whether the included children were screened for SNHL at birth was not available.

The children with HL were seen in these audiology clinics, usually either because of referrals from the screening programmes or referrals from general practitioners (GPs) or ear, nose, and throat physicians. Children with SNHL/MHL are followed regularly in the audiology clinics (at least every 6-12 months), whereas children with conductive hearing loss (CHL) are usually referred back to the physicians for further management.

Types of HL were defined as follows: SNHL was defined as abnormal air and bone conduction thresholds (>20 dB HL) without an air-bone gap (≤10 dB HL); MHL was defined as an abnormal air conduction and bone conduction threshold with an air-bone gap (>10 dB HL); and CHL was defined as abnormal air-conduction thresholds (>20 dB HL) with normal bone-conduction thresholds (≤20 dB HL) (Katz *et al.*, 2015). Because MHL has a sensorineural element, which is irreversible even if the conductive element has been treated/resolved, children with MHL were included in the SNHL group for the purposes of this analysis. Additionally, for the purpose of the analysis, children with CHL were merged with children with normal hearing in one group (the reference group).

In cases for which pure tone audiometry was used, the degree of HL was calculated by averaging the thresholds of the frequencies 0.5, 1, and 2 kHz. In cases for which auditory brainstem responses were measured, the air-conduction threshold of wave V was used to determine the degree of HL. This was the same approach used by audiologists working at the participating hospitals.

Based on the British Society of Audiology's classification, the degrees were defined as follows: mild HL ranging from 20-40 dB HL, moderate HL ranging from 41-70 dB HL, severe HL ranging from 71-95 dB HL, and profound HL >95 dB HL (British Society of Audiology, 2011). For bilateral HL, the thresholds of the better ear were used, whereas for unilateral HL, the thresholds of the worse ear were used.

3.3.2 Study 2

In this cross-sectional survey, carers of children aged 0-12 years who attended the audiology clinics in four hospitals/centres in Riyadh (three of which were the same hospitals included in Study 1) during a 3-month period (September to November 2016) were recruited (n=190).

Audiologists in each clinic informed the carers about the study, provided them with the information sheet, and obtained written consent to participate. Subjects completed a questionnaire (see Appendix A), which was developed based on the risk factors of SNHL listed by the joint committee on infant hearing (JCIH) (American Academy of Pediatrics, 2007). Questions about specific factors thought to have an association with SNHL in Saudi children, such as consanguinity and family size (Zakzouk *et al.*, 1995), were also included. The final question asked about the child's current hearing test results was answered by the audiologists from the child's medical records to provide the objective diagnosis.

Ethics approvals for Studies 1 and 2 were provided by the ethics committees at the University of Southampton in the UK (identification numbers 20827 and 23832) and the participating hospitals/centres in the KSA.

3.3.3 Data analysis

Statistical analysis was performed using IBM SPSS version 24. Summary statistics were described using means, standard deviations (SD), and percentages. Univariable associations between categorical variables were tested using the Chi-square test. The association between each potential risk factor as an indicator and the presence of SNHL as an outcome was assessed. However, for cells that had values less than 5, Fisher's Exact Test was used. Univariable logistic regression analysis was performed for each risk factor of SNHL to estimate the magnitude of the association. Multivariable logistic regression modelling was performed including factors that showed statistically significant association with SNHL in the univariable analysis. SNHL as an outcome was compared against a reference category composed of a combination of CHL and no diagnosis of HL. A statistical significance cut-off value of p=0.05 was used for all the analyses.

3.4 Results

3.4.1 Study 1

The electronic medical records of 1224 children were reviewed. Of those, 1166 were included in the study, and 58 were excluded because of inconclusive hearing test results. The children's ages ranged between 0.1-10.4 years (mean= 3.8 years, SD= 2.7 years). HL was identified in 46% (n= 533) of the sample. Of those, 64% (n= 340) were identified with SNHL and 36% (n=193) were identified with CHL. Descriptive data including information on the numbers and percentages of children identified with different types of HL are provided in Table 3.1.

The mean AOI of SNHL was 3.2 years (SD= 2.5 years, range: 0.1-10 years) and the median was 2.5 years (Interquartile range (IQR)= 1.1, 5). Children were distributed into four groups according to the AOI of SNHL: the first year of age (0-1 year), before school entry (1.1-3.0 years), during kindergarten ages (3.1-6.0 years), and after starting primary education (6+ years). Figure 3.1 shows the number of children identified with different degrees of bilateral and unilateral SNHL in different age groups.

Table 3.1 Numbers and percentages of children identified with different types of HL (Study 1)

	Number of children	Percentage of children	Mean age (SD) in years	Median age (IQR) in years	Mean age of identification (SD) in years	Median age of identification (IQR) in years
Male	703	60	3.8 (2.6)	3.4 (1.6-5.9)	-	-
Female	463	40	3.7 (2.7)	3.1 (1.6-5.8)	-	-
No HL	633	54	3.2 (2.5)	3.0 (1.0-5.0)	-	-
HL (SNHL* or CHL)	533	46	4.4 (2.7)	3.9 (2.2-6.6)	-	-
SNHL*	340	29	4.8 (2.8)	4.9 (2.5-7.0)	3.2 (2.5)	2.5 (1.1-5.0)
CHL	193	17	3.9 (2.5)	3.2 (2.0-5.9)	-	-
Total number of children	1166	-	3.8 (2.7)	3.2 (1.6-5.9)	-	-

*SNHL includes children with SNHL and children with MHL

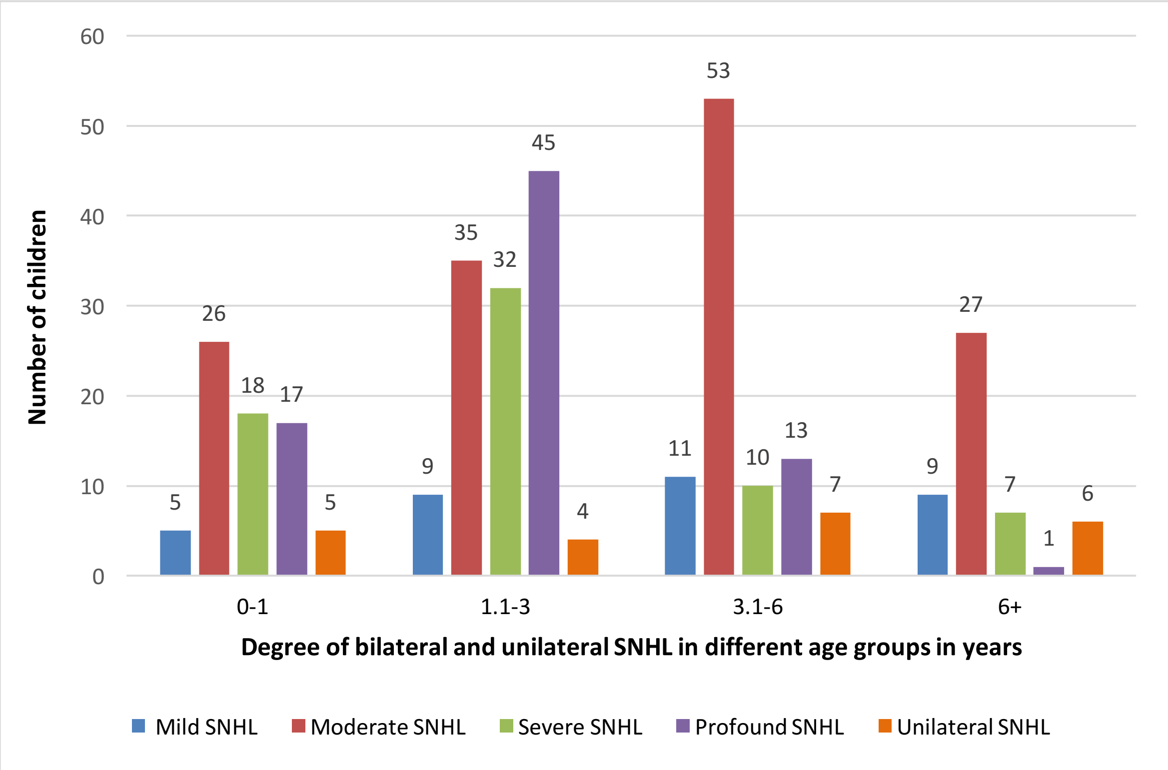


Figure 3.1 Number of children identified with different degrees of bilateral and unilateral SNHL in different age groups (Study 1)

A considerably high percentage of children with bilateral severe/profound degrees of SNHL was not identified until after their first year of age. Only 35 (24%) of bilateral severe/profound cases were identified within their first year. The remaining 108 (76%) were identified at later ages with 31 (22%) children not identified until after their 3rd birthday (after starting their schooling). Other degrees of SNHL showed similar results with only a few (n= 31, 18%) of the children identified with bilateral mild/moderate SNHL during their first year of age, whereas more than half of the children with bilateral mild/moderate SNHL were not identified until after school age (3+ years) (n= 100, 57%). Unilateral SNHL was identified in 22 (7%) of the children identified with SNHL in the sample. Of these children, 13 (59%) were identified after school age (3+ years). One child with profound unilateral SNHL was not identified until 9 years of age. Another 2 children with severe unilateral SNHL were not identified until the age of 9.4 and 10 years.

Several risk factors of HL (as defined by JCIH, 2007) were explored in the sample. Of those, 7 were found to be significantly more common in children identified with HL (any type) compared to children with normal hearing (no diagnosis with HL). The Chi-square test showed statistically significant difference between children who were identified with HL (any type) and children who had no diagnosis with HL in the following factors: parental consanguinity (p= 0.005), positive family history of HL (p<0.0001), admission to NICU for more than 5 days (p= 0.001), having any syndrome (p< 0.0001), treatment with chemotherapy (p= 0.003), having craniofacial anomalies (p <0.0001), and history of brain pathology/tumour (p= 0.002) (Table 3.2).

The presence of several predictors of SNHL (as defined by JCIH, 2007) in the children in the sample and the association between these predictors and the presence of SNHL were explored. The children with CHL and those with no diagnosis of HL were merged into one group and compared to children with SNHL using univariable logistic regression analysis. Table 3.2 provides information on the association between each potential risk factor as an indicator and the presence of SNHL as an outcome. Associations (p<0.05) were found between the presence of SNHL and parental consanguinity (p= 0.001), positive family history of SNHL (p< 0.0001), NICU admission for more than 5 days (p< 0.0001), treatment with chemotherapy (p= 0.001), and the presence of a diagnosis of brain pathology/tumour (p= 0.001). The multivariable regression model, which takes into account all of the above statistically significant factors, did not alter these results, except that the NICU admission predictor ceased to be associated with the presence of SNHL (p= 0.09). Consanguineous marriages were found in 62% of the parents of children with complete data on this variable. The children of consanguineous marriages were 1.7 times more likely to have SNHL than the children of non-consanguineous marriages (95% confidence intervals (CI): 1.2 to 2.4, p= 0.005). Moreover, children with positive family history of SNHL, history of chemotherapy, and/or history of brain pathology/tumour were 4, 5 and 7 times more likely to have SNHL than children free of those factors (Table 3.2).

Table 3.2 Characteristics of children with HL and the association of the presence of SNHL/CHL with different risk factors (Study 1)

	No HL **n (%)**	**SNHL** **n (%)**	**CHL** **n (%)**	**Number of children**[a]	**Chi-square**[b] ***p*-value**	**Unavailable regression odd ratio of SNHL**[c] **(95% CI)** ***p*-value**	**Multivariable odd ratio of SNHL**[d] **(95% CI)** ***p*-value**
Parental consanguinity	267 (58%)	167 (71%)	80 (61%)	833	0.005	1.7 (1.2 – 2.3) 0.001	1.7 (1.2 – 2.4) 0.005
Family history of SNHL	44 (7%)	78 (27%)	9 (6%)	1045	<0.0001	4.8 (3.3 – 7.0) < 0.0001	4.2 (2.7 – 6.5) <0.0001
Admission to NICU >5 days	213 (36%)	72 (24%)	67 (37%)	1071	0.001	0.5 (0.4 – 0.8) < 0.0001	0.7 (0.5 – 1.1) 0.09
Having any syndrome	62 (10%)	42 (12%)	46 (24%)	1166	<0.0001	0.9 (0.6 – 1.4) 0.7	-
Head trauma	2 (0.3%)	2 (0.6%)	2 (1%)	1166	0.32	1.2 (0.2 – 6.7) 0.8	-

	No HL n (%)	SNHL n (%)	CHL n (%)	Number of children[a]	Chi-square[b] *p*-value	Unavailable regression odd ratio of SNHL[c] (95% CI) *p*-value	Multivariable odd ratio of SNHL[d] (95% CI) *p*-value
Renal disease	11 (2%)	8 (2%)	7 (4%)	1166	0.29	1.1 (0.5 – 2.5) 0.9	-
Hyperbilirubinemia	8 (1%)	5 (2%)	2 (1%)	1166	0.94	1.2 (0.4 – 3.6) 0.7	-
Treatment with chemotherapy	15 (2%)	20 (6%)	2 (1%)	1165	0.003	3.0 (1.5 – 5.7) 0.001	4.9 (1.7 – 14.7) 0.004
Craniofacial anomalies	18 (3%)	19 (6%)	34 (18%)	1166	<0.0001	0.9 (0.5 – 1.5) 0.7	-
Brain pathology/tumour	10 (2%)	19 (6%)	4 (2%)	1166	0.002	3.4 (1.7 – 6.9) 0.001	7.1 (2.4 – 21.0) <0.0001
Hypoxia	12 (2%)	8 (2%)	4 (2%)	1166	0.85	1.2 (0.5 – 2.9) 0.7	-

	No HL **n (%)**	**SNHL** **n (%)**	**CHL** **n (%)**	**Number of children[a]**	**Chi-square[b]** ***p*-value**	**Unavailable regression odd ratio of SNHL[c] (95% CI)** ***p*-value**	**Multivariable odd ratio of SNHL[d]** **(95% CI)** ***p*-value**
Postnatal infectious disease	23 (4%)	13 (4%)	5 (3%)	1166	0.73	1.1 (0.6 – 2.2) 0.7	-
Neurological disorder	19 (3%)	15 (4%)	3 (2%)	1166	0.21	1.7 (0.9 – 3.3) 0.1	-

[a] Number of children excluding the missing data
[b] HL compared against a reference category composed of children who have no diagnosis with HL
[c] SNHL/MHL compared against a reference category composed of a combination of CHL and no diagnosis with HL in univariable logistic regression
[d] Multivariable logistic regression for the variables which had a statistically significant association with SNHL in the univariable analysis, consanguinity, family history of SNHL, Admission to NICU >5 days, chemotherapy and brain pathology/tumour

3.4.2 Study 2

Of the 190 carers of children who were surveyed, 174 of their questionnaires were included and 16 were excluded due to incompletion, giving a response rate of 92%. Children's ages ranged between 0.1-12.0 years (mean= 5.5 years, SD= 3.5 years). HL was identified in 56% (n=97) of the children. Of those, 82% (n=80) have SNHL and 18% (n=17) have CHL. Table 3.3 provides descriptive data of the sample including information on the number and percentages of children identified with different types of HL. As can be seen in Table 3.3, just below half of the children (46%) were identified with SNHL. Of those, 42 (53%) have bilateral severe/profound degrees of SNHL with 2 children identified after primary school age (6+ years). Unilateral SNHL accounted for 10% (n= 8) of the children identified with SNHL.

The mean AOI of SNHL was 3.1 years (SD = 2.6 years, range: 0.1-9.0 years) with a median of 2 years (IQR= 1, 4.8). Carers' responses showed that almost one-third of the children with SNHL of any degree were identified during the first year of age (31%, n= 25), whereas the majority were identified at later ages. Between 1 and 3 years, 33% (n= 26) were identified with SNHL, another 20% (n= 16) were identified between 3 and 6 years, whereas the remaining 16% (n= 13) were not identified until after their 6th birthday, which is after they started their primary schooling. Two children with unilateral SNHL were identified beyond 6 years (both identified at 8 years old). Figure 3.2 shows the number of children identified with different degrees of bilateral and unilateral SNHL.

Carers were asked about the presence of various risk factors of SNHL in their children. The Chi-square test showed statistically positive associations between the presence of HL (any type) and parental consanguinity (p= 0.001), parental concern about the child's hearing (p< 0.0001), having any syndromes (p=0.002), and having craniofacial anomalies (p< 0.0001) (Table 3.4). The association of various risk factors with the presence of SNHL was explored. Both univariable and multivariable logistic regression analyses of these variables showed that parental consanguinity and parental concern about the child's hearing are possible predictors for SNHL in children (p< 0.0001) in comparison to a reference group consisting of children with CHL and children who have no diagnosis with HL (Table 3.4). It was found that 57% of the children in Study 2 have consanguineous parents. Children of related parents appeared to be 4.5 (95% CI: 2.1 to 9.5) times more likely to have SNHL than children of unrelated parents. The identification of SNHL in children whose parents had concerns about their hearing was 13.5 (95% CI: 5.5 to 33.0) times more than other children.

Table 3.3 Numbers and percentages of children identified with different types of HL (Study 2)

	Number of children	Percentage of children	Mean age (SD) in years	Median age (IQR) in years	Mean age of identification (SD) in years	Median age of identification (IQR) in years
Male	102	59	5.4 (3.6)	5.0 (2.0-9.0)	-	-
Female	72	41	5.6 (3.3)	6.0 (2.7-8.0)	-	-
No HL	77	44	5.2 (3.6)	5.0 (1.6-8.0)	-	-
HL (SNHL* or CHL)	97	56	5.7 (3.4)	5.6 (3.0-9.0)	-	-
SNHL*	80	46	5.4 (3.6)	5.0 (2.6-8.0)	3.1 (2.6)	2.0 (1.0-4.8)
CHL	17	10	6.7 (3.3)	6.0 (4.0-10.0)	-	-
Total number of children	174	-	5.5 (3.5)	5.2 (2.6-8.0)	-	-

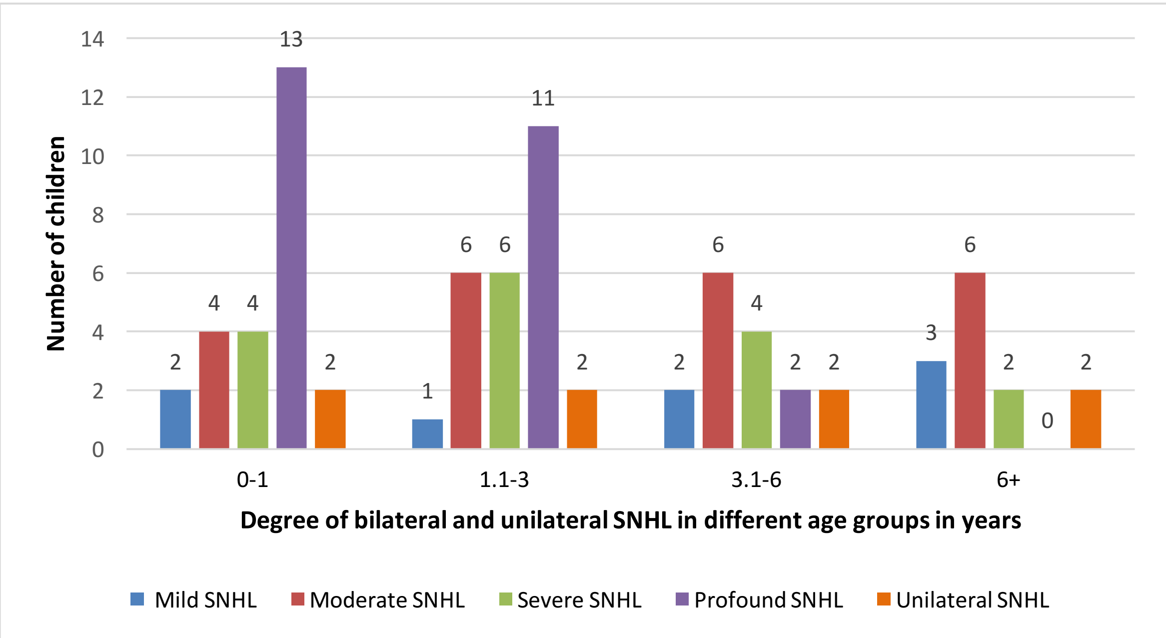


Figure 3.2 Number of children identified with different degrees of bilateral and unilateral SNHL in different age groups (Study 2)

Table 3.4 Characteristics of children with HL and the association of the presence of SNHL/CHL with different risk factors (Study 2)

	No HL N (%)	SNHL N (%)	CHL N (%)	Number of children[a]	Chi-square[b] *p*-value	Univariable regression odd ratio of SNHL[c] (95% CI) *p*-value	Multivariable odd ratio of SNHL[d] (95% CI) *p*-value
Maternity infectious diseases	1 (1%)	0	1 (6%)	174	0.09	-	-
Low birth weight <2 Kg	13 (17%)	8 (10%)	2 (12%)	174	0.48	1.7 (0.7 – 4.3) 0.3	-
Hypoxia	10 (13%)	8 (10%)	1 (6%)	174	0.77	1.2 (0.5 – 3.1) 0.7	-
Hyperbilirubinemia	19 (25%)	16 (20%)	4 (24%)	174	0.77	1.3 (0.6 – 2.7) 0.5	-
Admission to NICU	22 (29%)	20 (25%)	3 (18%)	174	0.69	0.9 (0.5 – 1.8) 0.8	-
Post-natal infectious disease	4 (5%)	4 (5%)	0	174	1	1.2 (0.3 – 4.9) 0.8	-

	No HL N (%)	SNHL N (%)	CHL N (%)	Number of children[a]	Chi-square[b] *p*-value	Univariable regression odd ratio of SNHL[c] (95% CI) *p*-value	Multivariable odd ratio of SNHL[d] (95% CI) *p*-value
Having any syndromes	7 (9%)	5 (6%)	6 (35%)	174	0.002	0.4 (0.1 – 1.2) 0.1	-
Craniofacial anomalies	0	7 (9%)	4 (24%)	174	<0.0001	2.2 (0.6 – 7.7) 0.2	-
Neurological diseases	5 (6%)	7 (9%)	0	174	0.57	1.7 (0.5 – 5.6) 0.4	-
Developmental delay	10 (13%)	13 (16%)	6 (35%)	174	0.08	0.9 (0.4 – 2.1) 0.9	-
Language delay	33 (43%)	33 (41%)	10 (59%)	174	0.42	0.8 (0.5 – 1.5) 0.6	-
Head trauma	7 (9%)	6 (8%)	1 (6%)	174	0.92	0.9 (0.3 – 2.6) 0.8	-
Treatment with chemotherapy	1 (1%)	4 (5%)	0	174	0.49	4.9 (0.5 – 44.7) 0.2	-

	No HL N (%)	SNHL N (%)	CHL N (%)	Number of children[a]	Chi-square[b] *p*-value	Univariable regression odd ratio of SNHL[c] (95% CI) *p*-value	Multivariable odd ratio of SNHL[d] (95% CI) *p*-value
Medication-induced HL	4 (5%)	10 (13%)	2 (12%)	172	0.20	2.2 (0.8 – 6.2) 0.2	-
Parental consanguinity	33 (43%)	58 (73%)	8 (47%)	173	0.001	3.3 (1.8 – 6.3) < 0.0001	4.5 (2.1 – 9.5) *p*<0.0001
Family history of SNHL	9 (12%)	16 (21%)	4 (24%)	170	0.24	1.6 (0.7 – 3.5) 0.3	-
Parent concern	28 (38%)	72 (90%)	13 (77%)	170	<0.0001	10.8 (4.6 – 24.9) <0.0001	13.5 (5.5 – 33.0) *p*<0.0001
Teacher concern	10 (14%)	7 (9%)	3 (19%)	168	0.43	0.6 (0.2 – 1.6) 0.3	-
Mother university education	43 (56%)	38 (48%)	8 (47%)	174	0.56	0.8 (0.4 – 1.4) 0.4	-
Father university education	45 (58%)	41 (51%)	9 (53%)	174	0.68	0.8 (0.4 – 1.4) 0.4	-

	No HL N (%)	SNHL N (%)	CHL N (%)	Number of children[a]	Chi-square[b] *p*-value	Univariable regression odd ratio of SNHL[c] (95% CI) *p*-value	Multivariable odd ratio of SNHL[d] (95% CI) *p*-value
Mother employment	24 (31%)	25 (31%)	4 (24%)	174	0.86	1.1 (0.7 – 2.1) 0.8	-
Smoking at home	26 (34%)	22 (28%)	6 (35%)	172	0.74	0.8 (0.4 – 1.5) 0.4	-
Low family monthly income (<10,000 SR)	44 (57%)	50 (63%)	14 (82%)	174	0.15	1.0 (0.6 – 1.9) 0.9	-

[a] Number of children excluding the missing data
b HL compared against a reference category composed of children who have no diagnosis with HL
[c] SNHL/MHL compared against a reference category composed of a combination of CHL and no diagnosis with HL in univariable logistic regression
[d] Multivariable logistic regression for the variables which had a statistically significant association with SNHL in the univariable analysis, consanguinity and parent concern

Carers of the children who were diagnosed with SNHL after the age of 3 years were asked their opinions about the reason for the seemingly late diagnosis (n = 29). Different reasons were given by carers, as can be seen in Figure 3.3.

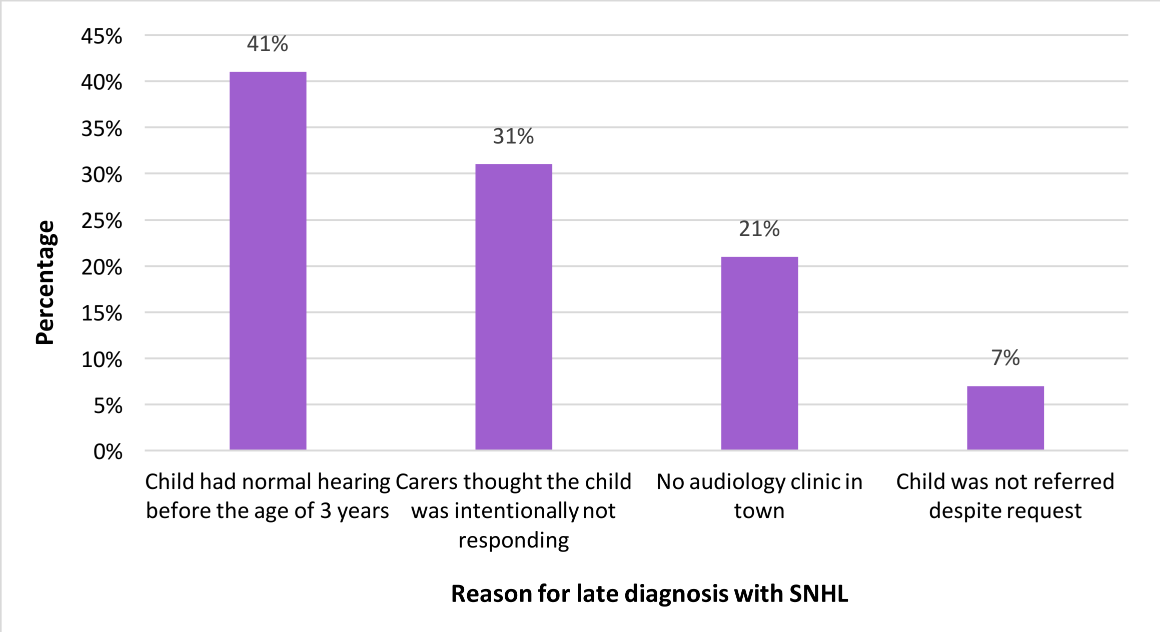


Figure 3.3 Carers' opinion about reasons for late diagnosis with SNHL

Carers of children who felt their children had developed SNHL after 3 years of age (n= 12) were asked to comment on what caused them to feel that way. Reasons, as reported by carers, are provided in Figure 3.4.

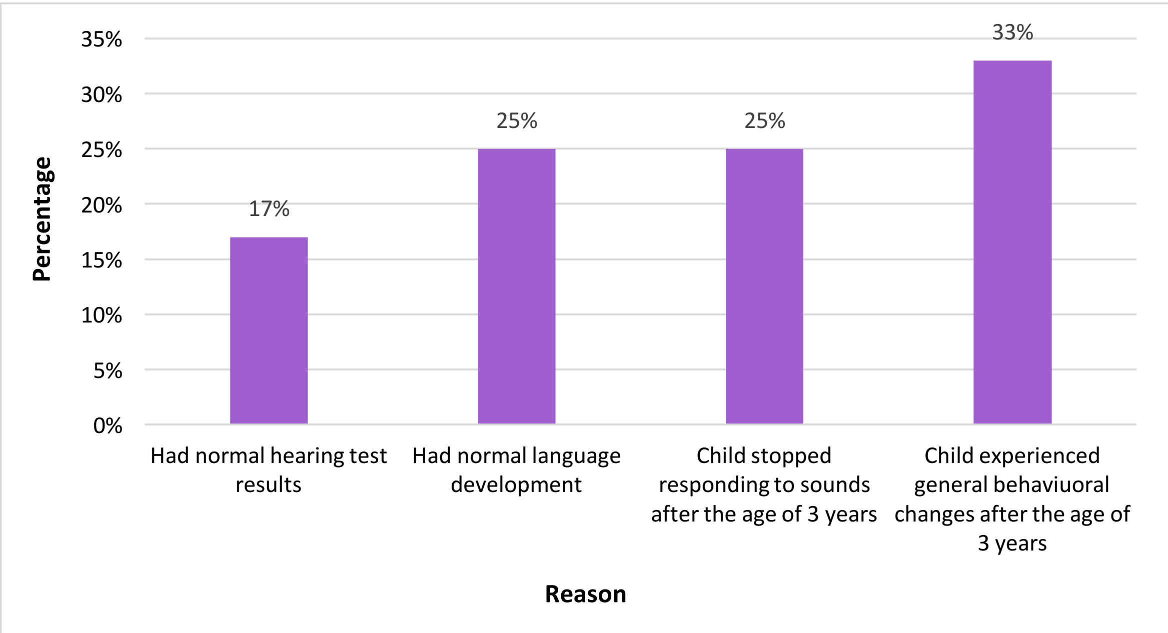


Figure 3.4 Reasons for carers to think the child's hearing was normal before the age of 3 years

When asked their opinion about an implementation of school entry hearing screening (SEHS) and whether they would allow their children to participate, 90% of carers agreed, 0.6% disagreed, 4.4% were unsure and needed further information about the screening process, and 5% deemed the question inapplicable to them since their children had already been identified with HL.

3.5 Discussion

Studies 1 and 2 aimed to provide initial data on the AOI of SNHL in Saudi children and to explore the current characteristics of children with SNHL to identify any potential risk factors that had not been identified earlier. The percentage of children identified with HL (any type) in both samples were considerably high (46% & 56% in studies 1 and 2 respectively) when compared to findings of other Saudi studies (Abolfotouh *et al.*, 1995; Al-Abduljawad and Zakzouk, 2003; Al-Rowaily *et al.*, 2012; Alharbi and Ahmed, 2015). This was not surprising since the samples of Studies 1 and 2 were attained from audiology clinics that most children visited because of hearing-related concerns.

The number of children identified with SNHL in both samples was more than those identified with CHL (Tables 3.1 and 3.3). A possible explanation is that most children with CHL are usually treated in primary care clinics, and only cases requiring further investigation are referred to specialised clinics. In addition, the referral period is usually long. Thus, CHL and its underlying causes might have been resolved during this period and children's hearing might have returned to normal limits by the time the child attended the audiology appointment. The following section discusses the findings related to the AOI of SNHL, followed by a section discussing the characteristics of children with SNHL.

3.5.1 Age of identification of SNHL in children in the KSA

The median AOI of childhood SNHL was similar for Studies 1 and 2: 2.5 years and 2 years, respectively. A widespread of the AOI in the samples was noticed: 75% of the median was almost 5 years in both studies, suggesting that some children reached school age before they were identified with SNHL. This is much higher than the AOI recommended by the JCIH (i.e. 3 months) (JCIH, 2007). In other HICs, such as the UK, the median AOI is 49 days (Wood *et al.*, 2015). Likewise, in the USA, the identification of childhood SNHL occurs on average before the age of 3 months (Dalzell *et al.*, 2000).

The findings of both studies revealed that high percentages of children with different degrees of SNHL were not identified until beyond the age of 3 years, including children with moderate and worse degrees of SNHL which are typically targeted by the UNHS (Public Health England, 2016). A possible explanation is that some children who visited the audiology clinic in the hospitals/centres, which have UNHS, were born in other hospitals and thus were not screened for SNHL at birth. Other explanations include being missed by the screening test, failing to follow-up after the first screening test, developing late-onset SNHL or acquiring SNHL. Regardless of the reason, having some children with bilateral severe/profound SNHL not identified until after school age is dramatic and serious.

It is acknowledged that the samples in Studies 1 and 2 were driven from a few audiology clinics located in two cities in the country. Thus, the findings might not be generalisable on a national level. However, the following clarification made the author confidently report that the present sample represents other regions of the country as well as Riyadh and Dammam. Patients who visit the audiology clinics included in the 2 studies (especially in Riyadh) come from different regions of the country because 2 of the included hospitals are main military and civilian medical cities that serve people who live around the country. Because of the lack of audiology clinics in some regions, eligible people who live outside of Riyadh are served at the audiology clinics located in Riyadh. For instance, the Security Forces Hospital in Riyadh (one of the included hospitals) is the main hospital that serves people who work for the different sectors of the Ministry of the Interior with tertiary health services. Eligible people might be living anywhere in the country where audiology clinics might not be available; for that reason, the audiology clinic at the main hospital in Riyadh accepts referrals from primary care units that follow the hospital and are located all around the country.

In an attempt to explain the high AOI of childhood SNHL in the present samples, it was difficult to tell if the identification was late or if the SNHL had developed or had been acquired at a late age. This is because the age of onset of SNHL was unknown; the only available information was the age at which SNHL was identified/confirmed. This leads to two explanations for this high AOI: the first is that the children in the present samples had congenital SNHL but had not been identified until older and the second is that the children might have had normal hearing at birth but developed/acquired SNHL at later ages.

The first explanation is possibly true because it is well known that most forms of hereditary SNHL are autosomal recessive in nature (80%) and would cause prelingual SNHL (Schrijver, 2004). A previous study reported that the prevalence of autosomal recessive SNHL in the KSA is 15.59% (Al-Qahtani *et al.*, 2010). The high practice of consanguinity in the country (Zakzouk *et al.*, 1993) is expected to cause autosomal recessive SNHL because consanguinity is known to have a high chance of pairing recessive alleles of SNHL (Shawky *et al.*, 2013). The findings of Studies 1 and 2 are consistent with this possibility of late identification because the majority of children identified with SNHL in the samples had consanguineous parents. The findings also showed that parental consanguinity is a potential predictor of childhood SNHL. Even in Study 2, where the sample size was considerably smaller, parental consanguinity was still shown as a potential predictor of childhood SNHL. It was difficult to tell if the late AOI (if it is assumed to be late) occurred because of the absence of UNHS, or even if it occurred with the implementation of UNHS, because the samples included children who might have been screened for SNHL at birth and children who were not. The exact information was not available.

The second explanation for the high AOI (appearance of late-onset SNHL or acquired SNHL) is also possible because late-onset SNHL, which is progressive in most cases (Schrijver, 2004), is also related to consanguinity. This was reported in the KSA where consanguinity was found in 81% of the cases with known progressive SNHL (Zakzouk *et al.*, 1995). Additionally, the possibility of acquiring SNHL at a late age because of environmental factors is acknowledged. Several Saudi studies reported that late-onset and acquired SNHL are prevalent in children in the KSA (al Muhaimeed, 1996; al Muhaimeed and Zakzouk, 1997; Almuneef *et al.*, 1998; Al-Noury, 2011). In Study 2, just below half (41%) of the carers of children diagnosed with SNHL after the age of 3 years reported that their children had a normal hearing before the age of 3 years and gave different justifications (Figure 3.4). These findings support the explanation that some children might have developed/acquired SNHL at a late age. Another fact that supports this explanation is that in other countries such as the UK, Australia and the USA, where the practice of consanguinity is not as common as in the KSA (Romeo and Bittles, 2014), the median AOI of SNHL prior to the implementation of UNHS was earlier than it is in the KSA (18.1, 18.5 and 24+ months in the UK, Australia and the USA respectively) (Harrison and Roush, 1996; Fortnum and Davis, 1997; Wake *et al.*, 2005). It appears that both explanations seem to apply at least in part to some children in the present samples.

Knowing that these findings were taken from Riyadh and Dammam, which host tertiary care hospitals where most audiology clinics are located, it is highly likely that the findings will be worse in other cities where there is less access to hearing healthcare. Hence, applying a hearing screening at early childhood in the KSA is highly recommended.

3.5.2 Risk factors of SNHL in children in the KSA

In the KSA, a prevalence of 56% of consanguineous marriages was reported (Zakzouk *et al.*, 1993). This is comparable to the findings of this research in which consanguinity accounted for 62% and 57% of the samples in Studies 1 and 2, respectively. These findings indicated that consanguineous marriages are still practised significantly in the country. Additionally, 71% and 73% of the children identified with SNHL in Studies 1 and 2, respectively, had consanguineous parents. This is close to what was reported in another Saudi study in which 81% of the children with SNHL had consanguineous parents (Zakzouk *et al.*, 1995). It is also similar to what was found in another Gulf country (Oman) with a reported prevalence of 70% of consanguinity among parents of children with SNHL (Al Khabori and Patton, 2008). The results of Studies 1 and 2 suggested that children who are products of consanguineous marriages seem to be 1.7 and 5 times more likely to have SNHL than children who are products of unrelated parents, respectively.

The high practice of consanguineous marriages in the country increases the possibility of having congenital and late-onset SNHL (Zakzouk *et al.*, 1995; Schrijver, 2004; Shawky *et al.*, 2013). Since preventing this practice is not possible due to cultural, social, political and economic factors that have historically favoured these marriages, periodic hearing screening for children is recommended. The findings showed that 90% of carers would allow their children to be screened for HL at schools if a hearing screening programme was available, which is encouraging. Additionally, public health campaigns and genetic counselling are highly recommended in order to increase societal awareness of the potential consequences of consanguinity, including SNHL.

Parental concern regarding children's hearing, which is recognised as a risk factor of SNHL in other countries (American Academy of Pediatrics, 2007), was identified for the first time as one of the potential predictors of SNHL in Saudi children. Apparently, parents may have a suspicion about their child's hearing, but the child still does not receive a hearing assessment early enough. This could be drawn back to several reasons, according to what was found in Study 2. First, the lack of audiology clinics in some regions of the country, which makes it difficult for people living in those regions to gain access to audiology services. The huge area and the concentration of the audiology clinics in the big cities requires people to travel for long distances, which costs them time and money. This is a major obstacle that requires great attention from all stakeholders to find the best solutions that would be suitable for the situation in the KSA. One possible solution is to provide a hearing screening tool that could be easily accessed and used outside the audiology clinics by individuals who have no experience in audiology (e.g. parents). This would help in screening the hearing of a child in case of suspicion without worrying about the difficulty accessing audiology clinics. The availability of such a test would be useful to parents who live away from primary care units and audiology clinics to screen the hearing of their child in case of suspicion and to seek help only if the child fails the test.

Second, the limited awareness of some GPs and carers of the importance of dealing with the suspicion of HL in children urgently needs improving. The limited awareness of carers appeared in the form of blaming the child for not responding whereas the limited awareness of GPs appeared in the form of delayed referrals to an audiology clinic regardless of the carers' request. Apparently, the awareness of GPs in the KSA about the seriousness of HL and its consequences if untreated or treated with delay is lower than that of the GPs in other HICs such as the UK. While the majority of children with HL in areas that do not apply SEHS in the UK was referred by GPs (Fortnum *et al.*, 2016), the limited awareness of some GPs in the KSA is one of the possible causes for the high AOI of SNHL in children. This highlights the need for raising the awareness of GPs to the importance of the immediate referral of children for hearing assessment in case of suspicion. It might be useful to provide a hearing screening tool that could be used by GPs in case of suspicion. This would help in identifying the affected children earlier and would help in reducing the number of referrals to the audiology clinics as only cases who failed the test would be referred.

Although it was reported in the literature that acquired SNHL resulting from postnatal infectious diseases is prevalent in children in the KSA (al Muhaimeed, 1996; al Muhaimeed and Zakzouk, 1997; Almuneef *et al.*, 1998), the findings of the current research showed no evidence of association between postnatal infectious diseases and the presence of SNHL. This might be explained by the small sample size in Study 2 and/or by a high commitment to vaccination programmes against childhood infectious diseases.

Another well-recognised cause of acquired SNHL is the use of chemotherapeutic agents and aminoglycoside antibiotics, which are known to cause ototoxicity (Madell and Flexer, 2008). In the KSA, 7.7% of children who received chemotherapeutic drugs showed SNHL after their first dose (Al-Noury, 2011). This is comparable to the findings of this research where 6% and 5% of children identified with SNHL in Studies 1 and 2 respectively were reported to have a history of chemotherapy. The findings revealed an association between chemotherapy and SNHL, suggesting that children treated with chemotherapy were 5 times more likely to develop SNHL than children who were not. Furthermore, the findings showed that having a history of brain pathologies/tumours is a possible predictor of SNHL in children. In the current sample, children with brain pathologies/tumours were 7 times more likely to have SNHL than those who did not. Again, the availability of an accessible and usable hearing screening tool would help in identifying SNHL in those children early after onset as it could be used to screen their hearing after each treatment session.

Except for parental consanguinity, factors that were suggested as potential predictors of SNHL in Study 1 were not in Study 2. However, a possible explanation of this is the small sample size in Study 2. Additionally, in Study 2, a carers' questionnaire was used to find the potential risk factors of SNHL. This method of data collection might not be accurate enough since carers might not recall the requested information precisely, which would affect the results. In contrast, in Study 1 information was derived from medical records, which is a more accurate method. An additional potential predictor of SNHL that was found in Study 2 is parental concern about the child's hearing abilities.

Other factors that were reported in the literature as posing a high risk of HL in children in the KSA include low socioeconomic status, mother's employment (Bafaqeeh *et al.*, 1994), large family size, low parental education and poor standard of hygiene (Ashoor, 1983). No association was found between those factors and SNHL in the present sample (Study 2). The small sample size could be a possible explanation of this. In addition, there could be a credibility issue in filling in the questionnaire, especially when answering sensitive questions such as questions about socioeconomic status, parental education, smoking etc.

The findings of Studies 1 and 2 raise the need for a potential hearing screening tool that could be used easily by non-audiologists. It also highlights the need for the application of periodic childhood hearing screening in the KSA. As mentioned previously, UNHS and school screening programmes have been launched recently in the country. It is expected for the UNHS to reduce the AOI of SNHL in children. However, there will still be a need for further hearing screening to take place to identify children who are not targeted by the UNHS (mild and unilateral cases), children who are missed by the UNHS and children who develop and/or acquire SNHL at later ages. Although the recently launched school screening programme is expected to identify those children, the time gap between the UNHS and school-based screening is large (6 years). Thus, by the time children with SNHL are identified through the school-based screening, they will be losing their sensitive period of brain plasticity, particularly those who were missed by the UNHS or who develop late-onset SNHL at early childhood. It is therefore considered in this research to develop a flexible hearing screening tool that could be used for formal and informal hearing screening to widen the possibilities of how a screen test could be implemented.

In addition, it might be useful to consider applying a targeted hearing screening to children who are at risk of SNHL. This could be done at schools for kindergarten children (3-5 years). The current PhD project therefore considered exploring the feasibility of implementing such a programme in the KSA. For that purpose, the risk factors of SNHL identified in Studies 1 and 2 were used to develop a checklist that could be used as a first-line hearing-screening tool to identify children who are at risk of SNHL (Appendix B). This could be followed by a hearing test using the developed hearing screening tool (see Chapter 4 for the development of the tool).

3.5.3 Implications to practice

It is the right of children with SNHL to have a chance of receiving intervention as early as possible especially in a country where the latest audiological rehabilitation technologies including cochlear implants are available (Al-Muhaimeed *et al.*, 2009). Reducing the AOI of SNHL in Saudi children is crucial and should be considered seriously and urgently. To achieve that, it is crucial to improve the audiology services provided and facilitate access to audiology clinics in the KSA. It is also recommended to provide a flexible hearing screening tool that could be used to screen children's hearing immediately in case of suspicion. Otherwise, late identification and intervention would negatively affect the child's language acquisition, speech perception, social and emotional wellbeing, and academic performance (World Health Organization, 2006; American Academy of Pediatrics, 2007). It would also affect these children's future chances for employment and work productivity, thus posing an additional economic cost to society (Honeycutt *et al.*, 2003), which could have otherwise been prevented or at least reduced.

3.6 Conclusions

Studies 1 and 2 aimed to provide initial data on the AOI of SNHL in Saudi children and to explore the current characteristics of children with SNHL to identify any potential risk factors that had not been identified earlier. Based on the findings of the two studies, it was concluded that:

- The AOI of SNHL in Saudi children is high with a high percentage of children not identified until after school age
- Late-onset SNHL or acquired SNHL, difficulty accessing audiology services in the country and limited awareness of SNHL and the consequences of untreated or lately addressed SNHL in children are possible reasons behind the high AOI
- Parental consanguinity, positive family history of SNHL, history of chemotherapy, history of brain pathologies/tumours, and parental concern about children's hearing seem to be potential predictors of SNHL in children in the KSA
- Reducing the AOI of SNHL in Saudi children is a priority to ensure that children with SNHL have a chance of receiving early interventions and thus acquire normal speech, language, social/emotional and behavioural skills
- It is recommended that public health campaigns and genetic counselling are established to increase societal awareness of the potential consequences of consanguinity, including SNHL

- It is recommended to develop an accessible and usable hearing screening tool to be used by GPs, parents, teachers etc. to screen the children's hearing in case of suspicion. This is believed to be a practical solution to compensate, to some extent, for the shortage of audiology clinics in the country
- Considering the implementation of a targeted preschool hearing screening in the KSA is recommended.

Chapter 4 Study 3: Developing, recording and equalising the intelligibility of the speech stimuli of the Paediatric Arabic Auditory Speech Test (PAAST)

4.1 Overview of the chapter

The second aim of this PhD project was to develop an Arabic speech-in-noise (SIN) test that could be used to screen for sensorineural hearing loss (SNHL), and that would add to the assessment services provided to children with SNHL in the Kingdom of Saudi Arabia (KSA). This chapter is divided into two parts. The first part outlines the first stage of the development of the test, called the Paediatric Arabic Auditory Speech Test (PAAST), in terms of selection, recording and processing of the speech materials. The second part outlines Study 3, where the intelligibility of the speech materials of the PAAST was equalised.

4.2 Introduction

It was found in Studies 1 and 2 that the age of identification (AOI) of SNHL in children in the KSA is high. Reasons behind this included the large-scale practice of consanguinity, which is related to the presence of both autosomal recessive SNHL and late-onset SNHL. While autosomal recessive SNHL is usually prelingual and could be identified by the UNHS, late-onset SNHL could not be identified by the UNHS as the symptoms do not manifest until a few years after birth. Other reasons for the high AOI of SNHL included the limited awareness of some general practitioners (GPs) and carers of the importance of the urgent dealing with suspicion of hearing loss (HL) in children and the difficulty in accessing audiology clinics in the country. Based on these findings, it was concluded that the availability of a flexible hearing screening tool that could be easily accessed and used by non-audiologists would be useful to identify children with SNHL early after onset.

Because the targeted children in this research are children with SNHL, the aim was to develop a hearing screening tool that is sensitive to SNHL. It was therefore decided to develop a SIN test. This is because, as discussed in Chapter 2, Sections 2.2.1, the effect of SNHL manifest largely in the reduced ability to understand speech in the presence of competing noise. To assess this ability, SIN tests have been used as they demonstrate the auditory communication abilities of the child by measuring children's ability to discriminate between speech stimuli in noise. The measurement, which is referred to as the speech reception threshold (SRT) in noise, identifies the signal-to-noise ratio (SNR) at which 50% of the speech stimuli are identified correctly.

The lack of Arabic SIN tests for children motivated the development of such a test in this PhD project. It is believed that the availability of the test would not only serve the aim of providing a hearing-screening tool but would also add to the diagnostic and rehabilitative audiological services provided to children in the KSA as it would help audiologists to identify the difficulties faced by the child and to set a proper management plan that would assist an individual child to overcome the auditory communication disabilities (the clinical importance of SIN tests was discussed in Chapter 2, Section 2.3.2).

The criteria of the developed test, called the Paediatric Arabic Auditory Speech test (PAAST), were selected to meet its requirements, which are: (1) screening tool, (2) suitability for young children; (3) sensitive to SNHL; (4) accessible; and (5) can be used easily by non-audiologists. Table 4.1 provides the criteria of the PAAST and the reasons for selecting each criterion. It is worth mentioning here that screening tools are required to be: (1) simple, quick and easy to interpret; (2) easily performed by paramedic staff or other personnel; (3) acceptable to the public; (4) accurate[2]; (5) reliable[3]; (6) sensitive[4]; and (7) specific[5] (Hall, 2003).

Table 4.1 Criteria of the PAAST

Criterion	Reason
Suitable for young children from the age of 3 years old	to allow for the identification of SNHL in children as early as possible
Use simple speech stimuli	to suit young children
Does not require higher cognitive function	to suit young children
Use simple instructions	to suit young children
Interesting procedure for young children	to attract young children
Require relatively short testing duration	to comply with screening tools criteria and to comply with the relatively short appointment times in clinics
Could be implemented in digital technology	to ease access to the test and to attract young children
Automated operation	to ease its use by non-audiologists

[2] Accurate tests give true measurement of the attribute under investigation (Kumar, 2011)
[3] Reliable tests provide almost similar results each time the test is repeated (Kumar, 2011)
[4] Sensitive tests identify individuals who have the disease as positive cases (Kumar, 2011)
[5] Specific tests identify individuals who do not have the disease as negative cases (Kumar, 2011)

It is crucial when developing an auditory speech perception test to ensure that the test assesses an individual's hearing rather than other abilities, such as language, cognitive and/or attention abilities. For instance, if some of the words of an auditory speech test were not familiar to a child and that child responded incorrectly to a test stimulus, one could not tell if the incorrect response resulted from the inability of the child to hear the word or from the unfamiliarity of the word to the child. Therefore, it is important when developing an auditory speech perception test that targets young children to ensure that the speech materials are within their vocabulary. It is also necessary to keep the test as simple and attractive as possible in order for the children to engage with it.

The development of the PAAST was inspired by the McCormick Toy Discrimination Test (MTT) because it was found that most of the criteria of the MTT were similar to the criteria of the PAAST. For instance, the MTT, which consists of 7 pairs of monosyllabic words (Table 4.2), evaluates the child's ability to discriminate between pairs of words that share the same vowel (McCormick, 1977). Having a pair of words with the same vowel makes it more difficult for children with SNHL to distinguish between the words because the recognition of speech usually depends more on vowels rather than consonants. This is because most of the power of speech is in the vowels; thus vowels are more detectible and intelligible at lower intensities (Gelfand, 2010). Also, the ability to identify vowels develops much earlier than the ability to identify consonants (Cole *et al.*, 1996; Johnson, 2000). Thus, having speech stimuli with this pairing feature ensures the assessment of the child's ability to discriminate correctly between speech sounds without relying only on identifying the vowels, which is considered easily identifiable and more important to recognise the message (Cole *et al.*, 1996).

Table 4.2 The speech materials of the McCormick Toy Discrimination Test (MTT)

Cup	Duck
Spoon	Shoe
Man	Lamb
Plate	Plane
Horse	Fork
Key	Tree
House	Cow

In addition, the MTT uses closed-set materials (simple words), which meet the criteria of simple speech stimuli that have little linguistic redundancy. Less redundant speech materials, such as words, are less predictable than more redundant materials such as sentences (Gelfand, 2009). The use of closed-set material is better for young children and children with limited vocabulary as it allows only for limited choices of responses (e.g. choose from a limited number of words/pictures). Hence, it does not require complex cognitive functions (Madell, 1998; Gelfand, 2009). Moreover, the developer of the MTT (Berry McCormick) selected test items that are not only familiar to young children, but that can also be illustrated by toys. According to him, this was because young children are interested more in toys, and the development of toy recognition comes before that of picture recognition (McCormick, 1977). However, for the development of the PAAST, it was decided to select words that can be illustrated in pictures rather than toys to ease the implementation of the PAAST in an iPad application. The implementation of the test in a tablet-based application would ease the accessibility to the test without requiring special equipment, which in turn would allow for a better spread of the test. It would also make the test more attractive and engaging to the children of the new generations who use electronic devices a lot, where they watch cartoons, play games, listen to music etc. Thus, providing the PAAST through iPads or other electronic devices is expected to be attractive and interesting to children.

The procedure of the MTT includes a familiarisation phase where the toys are presented to the child and he/she is asked to name them or point to the named toy to ensure that the child is familiar with the words. If a word is found to be unfamiliar to the child, the pair is excluded from the test. The flexibility of the test in terms of excluding the unfamiliar words makes it appropriately usable with most young children. Additionally, the MTT uses a carrier phrase ("Where is the…?", "Point to the…" or "Show me the…"), which helps in getting the child's attention in case the child becomes distracted during the test. Most importantly, the test does not require a verbal response from the child; pointing to the heard word is enough, so it would work better with shy children and/or children who have good speech perception but poor speech production.

In general, the MTT uses simple speech materials and a simple procedure, it does not require high cognitive functions such as memory, recalling etc., and it showed good reliability in assessing the functionality of the hearing system in young children (Ousey *et al.*, 1989; Palmer *et al.*, 1991; Summerfield *et al.*, 1994; Lovett *et al.*, 2013). Based on this, it was decided to develop a similar test for Arabic-speaking children. However, the PAAST is not a typical translation for the MTT, but it follows the same principle.

Prior to describing the process of selecting the speech materials of the PAAST, it was important to provide an overview of some properties of the Arabic language. This would help the reader to understand the reason behind the selection criteria.

4.2.1 Arabic language

The Arabic language has two forms, the colloquial (spoken) form and the standard form (Saiegh-Haddad and Joshi, 2014). Colloquial Arabic has different dialects that are specific to certain geographical regions and demographical backgrounds. It is an informal language that is used for daily communication and has no specific writing style. On the other hand, standard Arabic is united and well understood by Arab people in all Arabic countries (Saiegh-Haddad and Joshi, 2014). It is mainly used for formal purposes such as in praying, teaching, formal writing, media etc. The two forms of the Arabic language differ in terms of the vocabulary, grammar and phonology. For example, in the KSA, the sound /q/ in the standard form in pronounced /g/ in the colloquial language.

In their first years of life, children in Arab countries usually learn and use the colloquial language because it is the spoken language used by their family members and people around them. They are exposed massively to standard Arabic when they first join schools where it is the language used. Although some TV cartoon shows are in standard Arabic, the main form of language that is used by young children is the colloquial (Amayreh, 2003). Therefore, for the purposes of developing the PAAST, the selected words were used in the colloquial form. To the best of the author's knowledge, no Arabic SIN test is available for children in colloquial Saudi Arabic.

The standard Arabic language consists of 28 consonants and 3 long vowels, /a, i, u/, and their short cognates, considered as short vowels (Amayrah, 2000). Consonants are described by three features: manner of articulation, place of articulation, and voicing. Tables 4.3, 4.4 and 4.5 show the features of Arabic consonants as reported by Sabir and Alsaeed (2014). Short vowels in Arabic have the same characteristics of the long vowels /a, i, u/ but with short pronunciation. Those short vowels are illustrated by symbols written either above or below the Arabic written letters (◌َ is the short /a/), (◌ُ is the short /u/ and (◌ِ is the short /i/) (Alotaibi, 2010). However, different Arabic colloquial languages have some more vowels that are used by people in certain geographical regions but not in others. For instance, in the KSA, Kuwait, United Arab Emirates, Jordan and Egypt, the following three long vowels are also used /ɑ, e, ɔ/ (Alotaibi, 2010). These vowels were considered when selecting the speech materials of the PAAST.

Table 4.3 Places of Arabic consonants

Place	**Speech sound**
Bilabial	b m w
Interdental	Θ ð ð̲
Labio-dental	f
Alveolar	t̲ l n r s z s̲
Denti- alveolar	t d
Pharyngeal	ʕ ħ
Palatal	d̲ ʒ ʃ j
Velar	K g
Uvular	q x̲ ɣ
Glottal	h ʔ

Table 4.4 Manners of Arabic consonants

Manner	**Speech sound**
Plosive	d̲ b t̲ k q t d ʔ g
Affricate	ʒ
Fricative	f Θ ð s̲ z s ʃ x̲ ɣ h ħ ð̲
Nasal	m n
Liquids	l r
Semi-vowel	j w

Table 4.5 Voicing of Arabic consonants

Voiced sounds	b d g ʒ l m n z ɣ r d̲ ʕ j w ð̲
Voiceless sounds	t k q f Θ ð s s̲ ʃ h ħ t̲ ʔ
Semi-voiced	x̲

4.3 Selecting, recording and processing the speech materials of the PAAST

4.3.1 Selecting the speech materials of the PAAST

The process of word selection went through several stages. First, the word criteria were set as follows:

- Nouns
- Monosyllabic words in the structure of CVC (consonant, vowel, consonant)
- Familiar to children as young as 3 years old
- All words together represent almost all manners and places of articulation of the Arabic consonants
- All words together include the three main vowels and some of the short vowels
- Easily pictured

During the process of word selection, the sounds /ð, ذ/ and /s̱, ص/ were avoided because it was reported in the literature that all consonants of the Arabic language are achieved and produced by the age of 2 years except for these two sounds (Amayrah, 2000). Although the PAAST does not require the child to produce the words verbally, it was preferred not to include these two sounds but instead to include words that have sounds with the same features as these two sounds when selecting the speech materials for the PAAST.

A set of simple monosyllabic nouns that are expected to be familiar to children as young as 3 years old was proposed and used to form six lists of monosyllabic Arabic words. Each list included 7 pairs of words; each pair shares the same vowel. The proposed lists included words that, in total, had the 6 long vowels mentioned earlier, /a, i, u, ɑ, e, ɔ/, and the 3 short vowels. The lists were emailed to audiologists from different regions of the KSA (because different Saudi Arabic dialects are used in different regions of the country) and to audiologists from other five Arab countries (Kuwait, United Arab Emirates, Jordan, Palestine and Egypt) who carefully reviewed the words in terms of their applicability to children in their regions/countries. The reason for involving audiologists from other Arab countries was to consider the applicability of the words of the PAAST to Arabic-speaking children not only in the KSA but also in other Arab countries due to the lack of SIN tests for Arabic-speaking children. Hence, the benefit of the PAAST would not be limited to children in the KSA but rather to children in other Arab countries.

Upon the audiologists' feedback, a list was prepared. Following this, the familiarity of the words included in the list was checked with 30 Arabic children aged 3-4 years old in different Arab countries (KSA, Kuwait and the United Arab Emirates). A sheet with pictures that represented the words was sent to the mothers of the children. The mothers were asked to name each picture and ask the child to point to the picture that represents the word. Mothers were asked not to repeat the word if the child failed to point to the correct picture but instead to consider it unfamiliar to the child. The mothers then sent the responses of their children to the author. Words that were 90-100% familiar to the children were then used to prepare a semi-final list of words that was used later in a pilot study (Table 4.6). The list included 7 pairs of words, 6 of which have long vowels and 1 has a short vowel.

Table 4.6 The semi-final word list of the PAAST

Pair	Arabic word	Phonetic transcription	English meaning
1	باب	/bab/	door
	ناس	/nas/	people
2	فيل	/fil/	elephant
	ديك	/dik/	hen
3	بيت	/bet/	house
	عين	/ʕen/	eye
4	ثوب	/θɔb/	Arabic men’s dress
	موز	/mɔz/	banana
5	نار	/nɑɽ/	fire
	فار	/fɑɽ/	mouse
6	شمس	/ʃʌms/	sun
	كلب	/kʌlb/	dog
7	دود	/dud/	worms
	نور	/nuɽ/	light

It can be seen in Table 4.6, that the pair /ʃʌms/ and /kʌlb/ is not typically in the CVC form but instead includes consonant clusters in the form of CVCC. The reasons for including this pair is because it was the most familiar pair that has a short vowel. Besides, consonant clusters in the form of CVCC are common in Arabic (Amayrah, 2000) and can be found in other Arabic speech reception tests which use monosyllabic words (Garadat *et al.*, 2017). In addition, in the MTT, the words were in different structures (Table 4.2), and this did not affect the accuracy of the test.

4.3.2 Recording the speech materials of the PAAST

Recordings took place at an anechoic chamber, which is a soundproof chamber, located at the Institute of Sound and Vibration Research, University of Southampton. The speaker was a native Arabic female from the central region of the KSA (the author). She was chosen because she speaks a "white dialect", which is not related to a specific region and is easily understandable by most Arabic children (at least in Gulf countries).

In the MTT, the following carrier phrases are used: "Show me the…", "Where is the…?" and "Point to the…" (McCormick, 1977). The use of carrier phrases helps in maintaining the same intonation when the speaker pronounces the target words, and in attracting the child's attention without having a substantive effect on the SRT (Gelfand, 2009). The carrier phrase "Where is the…?", was used for the PAAST because the colloquial form of the word "where" is similar for most Arabic people who speak different Arabic dialects whereas different colloquial words are used for "show me" and "point", which would make it difficult to generalise the use of the test with children speaking different Arabic dialects. Additionally, the phrase "Point to the…" requires the use of pronouns, which differ, as a word, between genders. The use of "Point to the…" requires having two versions of the recordings, one for boys and one for girls. Therefore, it was decided to use the phrase "Where is the…?" as the carrier phrase in the PAAST.

4.3.2.1 Procedure

The speaker sat on a comfortable chair in the anechoic chamber. A sound level meter (SLM) from Brüel and Kjær (type 2230, serial number 1033351) was used as the recording device. It was positioned on a tripod to keep it stabilised at the level of the speaker's mouth (Figure 4.1). The distance from the microphone of the SLM to the speaker's mouth was 75 cm, which is the same distance used by Semeraro (2015) to record the coordinate response measure (CRM) SIN test. The SLM was connected to an RME BabyFace sound card (serial number 23601885), which converts the analogue signals to digital signals. The sound card, in turn, was connected to a MacBook Pro laptop (UOS 11991) with an Adobe Audition CS6 software, which was used to process the words and save the files.

Figure 4.1 Recording the speech stimuli of the PAAST

The speaker pronounced the words and the carrier phrase separately with a short silence after each of them to allow for editing later. Natural intonation with approximately the same loudness level and vocal effort was kept during the recording. The words were recorded through the SLM to the sound card to the Adobe Audition software on the laptop using a sampling frequency of 44.100 Hz with 16-bit of amplitude resolution. Each word was recorded 9 times, over a period of 2 days (5 recordings on the first day and 4 on the second) to allow for selecting the best recording. The digitised signals were stored in separate files for later editing.

When listening to the recordings, a distortion in the background was heard, thus these recordings were discarded. After changing some equipment and trying different microphone-to-mouth distances, the words were re-recorded by placing the microphone of the SLM at 15 cm from the speaker's mouth. The repeated recordings were then judged in terms of the quality, clarity, intelligibility and natural presentation by 4 native Arabic experienced audiologists and another 10 Arabic-speaking individuals (including 6 children). It was agreed that the words sounded clear and natural, therefore they were used for the final version of the test. Spectrograms of the recorded words are provided in Appendix C.

4.3.3 Word processing and noise generation

All words were band-pass filtered by 50-20,000 Hz using a 3rd order zero-phase filter to improve the quality of the words and avoid low-frequency artefacts. Additionally, unnecessary air/silence before or after the words were removed. This was important to synchronise the beginning of the words to be able to control the length of the gap between the word and the carrier phrase. Afterwards, very short raised-cosine onset and offset ramps were used either side of the word to ensure the sound file started and ended with amplitude values of zero. The recordings were edited, the root mean square (RMS) amplitude of each word was calculated, the words were scaled to the same RMS level and then each word was saved separately as .wav files (sound files).

To generate the noise, all words were concatenated in one audio file and a steady-state noise was generated based on the speech spectrum using a MATLAB code (see Appendix D). The resulting noise is a speech-shaped spectrum noise; that is a white noise filtered with the long-term average speech spectrum (LTASS) of the recorded speech. Figure 4.2 plots the power spectrum density of the target words (concatenated, no gaps) and the speech spectrum noise. The RMS level of the generated speech-shaped spectrum noise was then fixed to the RMS level of the concatenated speech.

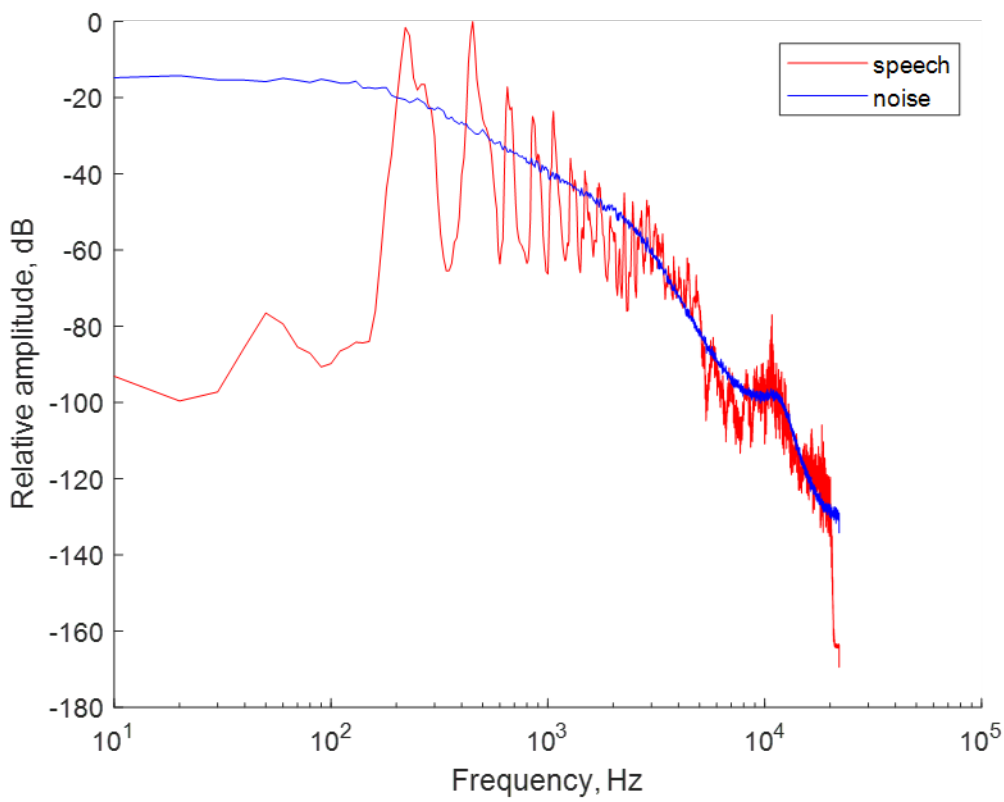


Figure 4.2 The power spectrum density of the target words and the speech spectrum noise

4.3.4 Summary

In the first part of this chapter, the first stage of developing the PAAST was completed. Seven pairs of monosyllabic words that share the same vowel and are familiar to young children were selected. Several senior audiologists revised the appropriateness of the selected words in terms of their acoustic properties and familiarity to young children in different Arab countries. The speech materials were then recorded and processed. After that, a noise that matches the LTASS of the speech stimuli was generated using a MATLAB code. At this point, the test was ready for the following stage, which is the intelligibility equalisation of the speech stimuli.

4.4 Study 3: Equalising the intelligibility of the speech materials of the PAAST in noise

In the first part of this chapter, the speech materials of the PAAST were selected, recorded and processed. The current section outlines the process of equalising the intelligibility of the speech materials (Study 3).

4.4.1 Introduction

The PAAST was developed to measure SRTs for Arabic-speaking children. Measuring SRTs is about measuring the subject's ability to *understand* the heard speech materials rather than the ability to detect the presence of a stimulus. In other words, it is a measure of the intelligibility of the received message. This makes the process more complicated than the measurement of pure tone thresholds or even speech detection thresholds. To measure a pure tone threshold, the tone is presented repeatedly at different stimulus levels in order to find the threshold. However, the case is different when measuring the intelligibility of speech stimuli, such as words, because if the same word was presented several times, the subject would identify the word when it is presented at high stimulus levels and then would respond correctly at the detection level where the word is not intelligible but rather detectable. To avoid this, a number of different words are selected and presented to the subject at the same stimulus level and the subject is required to distinguish between them. By doing this, it is ensured that the threshold of the intelligibility, not the detectability is measured. However, to guarantee accurate measurement, it is necessary to ensure that all the presented words are almost equally intelligible at a given stimulus level. Therefore, it is essential to assess and equalise the intelligibility of the words before using them for SRT measurements.

Equalising the intelligibility of the speech stimuli increases the test accuracy in threshold estimation when it is implemented in an adaptive procedure (see Chapter 5, Section 5.3.1 for further information about the use of adaptive procedures in threshold estimation). Otherwise, if a highly intelligible (too easy to identify) word is included in a set of speech-test stimuli, this particular word would achieve correct responses at low stimulus levels and thus would lead to estimate a threshold that is better than the true threshold and vice versa.

It is important to note here that the intelligibility of the speech stimuli is not only affected by the stimulus level (loudness of the stimuli), but other factors also contribute to it. For instance, the phonetic structure of the speech stimulus is one factor that would affect its intelligibility because different speech sounds have different sound energies (Gelfand, 2010). Thus, it is expected that at a given stimulus level, words that consist of sounds with more energy would be more intelligible than words with less powered sounds. Hence, words that are equally loud are not necessarily equally intelligible. Therefore, it is important to equalise the intelligibility of the words at a given stimulus level.

The presence of background noise is another factor that would negatively affect the speech intelligibility because the frequency components of the speech become more difficult to resolve in environments with competing signals. Although the ear acts as a filter to the received message where it passes the speech signal and removes a great deal of the noise, there is still some amount of noise passing through. This amount of noise is believed to affect the intelligibility threshold (Moore, 2003). Therefore, higher stimulus levels are required to understand speech in the presence of background noise.

It appears that it is essential to control other factors that would affect the intelligibility of the speech stimuli to ensure that at a given stimulus level, all speech stimuli show similar intelligibility. Therefore, it was decided to assess and equalise the intelligibility of the words of the PAAST in the presence of stationary speech-shaped noise (the type of noise used in the test).

To assess the intelligibility of the words, the psychometric function (PF) of each word is assessed. PFs describe the relationship between an observed behavioural performance on a psychophysical task and a given stimulus. They represent how behaviour changes in response to changes in the stimuli (Gescheider, 1997; Kingdom and Prins, 2010). For speech intelligibility in noise, PFs describe the relationship between the SNR and the proportion of correct responses and represent it with an S-shaped sigmoid curve (Figure 4.3) (see Appendix E for further information about PFs).

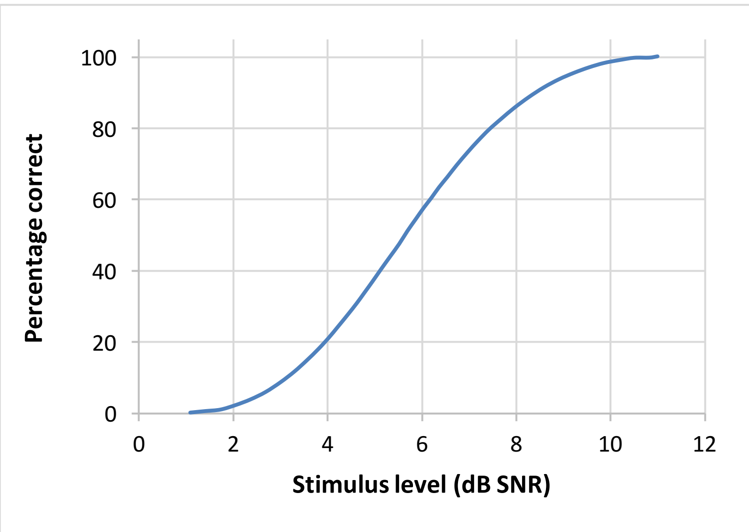


Figure 4.3 Graphical representation of a psychometric function

After fitting the PFs, two main features of the PF curve are assessed. These are the location, which represents where the curve is positioned in the x-axis (Gescheider, 1997; Kingdom and Prins, 2010), and the slope, which demonstrates the maximum rate at which the response changes as a result of a change in the stimulus level (MacPherson and Akeroyd, 2014). By assessing the locations of each word and calculating the average location, words that deviate from the average location are identified. Once they are identified, they are adjusted to equalise the intelligibility across all words (see the following section for more details about the equalisation process). Finally, the intelligibility of the words after adjustment is assessed to ensure that the adjustment was successful in equalising the intelligibility across the words.

4.4.1.1 Intelligibility equalisation

The main features of the PFs that are assessed to judge the intelligibility of the speech stimuli are the locations and slopes. The location is the only feature of the speech stimulus that can be manipulated to enhance homogeneity. Therefore, the location of each speech stimulus is assessed and compared to the mean location to check for any possible heterogeneity between the words. If heterogeneity (i.e. differences between the location of any word and the mean location) is noticed, several options are available to be done to enhance the homogeneity between the words. First, make no changes to the locations of all words. This option is unlikely to be used because it is expected that at least some speech stimuli will show largely deviated locations from the mean location and thus will require manipulation.

Second, the exclusion of all the words that show deviation from the mean location. It is not always practical to follow this option because it is expected for the majority of the words to show at least some deviation from the mean location and thus the exclusion would include large numbers of the speech stimuli. This option was followed by Ozimek *et al.* (2009), who equalised the intelligibility of the polish triple-digit test, where all the digit-triplets that deviated from the mean by ±1.5 dB were excluded (57 digit triplets were excluded). This was not problematic for them as the initial number of digit-triplets was 160 and thus even after the exclusion of the deviated stimuli, they still had enough speech stimuli (103).

Third, adjust the locations of the deviated words. This is usually done either by increasing or decreasing the amplitude of the word, which causes the PF to shift left or right to line it up with the average location (Figure 4.4). For example, if the average location is 15 dB and a word has a location of 11 dB, then the amplitude of that word should be decreased by 4 dB to line it up with the average location. This option is commonly used to equalise the intelligibility of SIN tests (e.g. Summerfield *et al.*, 1994; Potgieter *et al.*, 2016; Semeraro *et al.*, 2017; Sheikh Rashid *et al.*, 2017). However, it is worth noting that the authors of the mentioned studies equalised the intelligibility of the speech stimuli to a pre-defined threshold point (SRTn) on the x-axis of the PF curve, rather than the location (which is the steepest point on the x-axis).

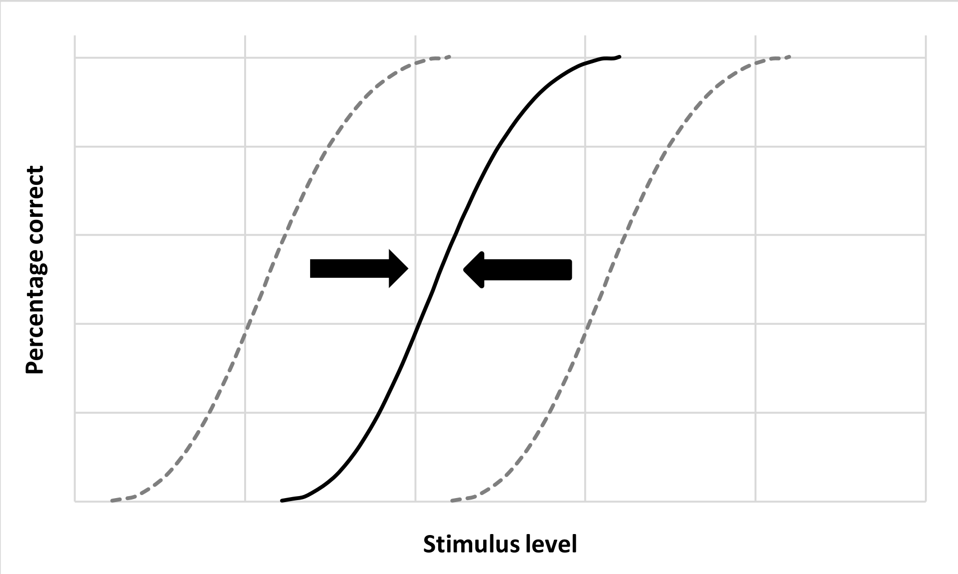


Figure 4.4 An example of adjusting the location of the deviated words to line up with the average location. The solid line represents a PF with an average location whereas the dashed lines represent PFs of deviated words

Fourth, use a combination of the three options by making no changes to the words that have locations that are close to the mean location, excluding the words that deviated largely from the mean location and adjusting the locations of the words that fell in between the two extremes (e.g. Smits *et al.*, 2004, Yuen *et al.*, 2019). Since no predefined definition is available about what location is considered close enough or far away from the mean location, it is left to the researcher to decide.

After adjustments, the PF of the adjusted speech stimulus is reassessed to verify that at a given stimulus level, all words give almost similar proportions of correct responses. Ideally, no difference should be there between the locations of all words after adjustments. However, because this is difficult to find in real practice, minimising the differences as much as possible is recommended.

It is worth noting that previous studies adjusted the speech stimuli to the threshold point on the x-axis rather than the location. In other words, they identified the level at which the SRT corresponds to 50% correct responses and then adjusted the speech stimuli to line up to that specific level (Plomp and Mimpen, 1979; Smits *et al.*, 2004; Ozimek *et al.*, 2009; Potgieter *et al.*, 2016; Semeraro *et al.*, 2017; Sheikh Rashid *et al.*, 2017; Yuen *et al.*, 2019). Other studies also adjusted the words to a threshold level, but they used different definitions of thresholds. For instance, Summerfield *et al.* (1994) and Lovett *et al.* (2013) adjusted their speech stimuli to line up with a threshold that corresponded to 70.7% of correct responses. However, to equalise the intelligibility of the words of the PAAST, the locations of the words, which represent the steepest point on the x-axis of the PF curve, rather than the threshold, were adjusted. This is because it was thought that equalising the locations of the words would allow the use of the speech stimuli of the PAAST for different purposes. For instance, the stimuli could be utilised in different adaptive procedures with no concerns about the procedure used in the intelligibility equalisation process. Additionally, the equalised speech stimuli could be used appropriately by researchers who are not interested in using an adaptive procedure but instead are interested in finding the threshold at a fixed intensity level. So, it could be said that equalising the locations is a more general approach.

In addition to assessing the locations of the speech stimuli, it is also important to assess their slopes to ensure that they show monotonicity and they are steep. A monotonic relationship means that increasing the stimulus level causes an increase in the performance level. Having steep slopes indicates the precision of the test score (Leensen *et al.*, 2011) because steep slopes indicate that a small change in the SNR elicits a change in the response. In other words, the steeper the slope, the greater the effect on intelligibility which results from a small change in the SNR. Conversely, for shallow slopes, a small change in the SNR would have little or no effect on the intelligibility and thus similar responses could be elicited from a range of SNRs. This would increase the potential for threshold estimation errors. So, it could be said that a steep slope reduces threshold estimation errors.

However, adjusting the amplitude of the speech stimuli does not affect the slopes. The factor that would help in having steep slopes is the use of appropriately distributed stimulus levels. Since there is no predefined definition in the literature about which slope is considered steep enough for the speech stimuli to be included in SIN tests, the stimulus levels should be selected carefully in advance to ensure performance responses that range from around a chance level to almost 100% correct responses. It is unnecessary to have many stimulus levels as long as the levels used are appropriately distributed to elicit responses from chance to 100% correct. To ensure steep slopes, selecting stimulus levels that result in large changes in the performance level with each change in the stimulus levels is recommended (Kingdom and Prins, 2010).

4.4.1.2 Aims

The aim of this study was to equalise the intelligibility of the words of the PAAST when presented in stationary speech-shaped noise. This aim was achieved by working through the following objectives:

- establish how similar the intelligibility of the words of the PAAST are when presented in a stationary speech-shaped noise by fitting the PFs of each word
- identify the words that deviated from the average location, calculate the amount of the adjustment needed and adjust the locations of the deviated words to equalise the intelligibility across the words
- ensure that the adjustments that were made enhanced the homogeneity between the words. In other words, ensure that at a given stimulus level, the performance level of all words is almost similar
- ensure that all words have relatively steep slopes.

4.4.2 Methods

There is no method to calculate the sample size for such studies where the aim is to gather enough data to obtain an accurate estimation of the PFs. Therefore, a sample size of 30 was selected to match or be larger than previous studies which assessed the PFs of speech stimuli (Plomp and Mimpen, 1979; Summerfield *et al.*, 1994; Smits *et al.*, 2013; Vaez *et al.*, 2014; Potgieter *et al.*, 2016; Semeraro *et al.*, 2017).

A total of 30 NH Arabic-speaking adults (18 females and 12 males) were included in the main study. Subjects' ages ranged between 18 and 54 years (mean = 30.2 years, SD = 8.8 years). Adults were tested because the subjects were required to attend 2 sessions of around 2 hours each. Children were not expected to tolerate attending such considerably long sessions. Besides, it is not uncommon to test adults to equalise the intelligibility of speech stimuli for SIN tests even when the test was developed for children (e.g. Summerfield et al., 1994).

Subjects were contacted personally and upon their willingness to participate, they were provided with an information sheet and a consent form to understand the details of the study and to confirm their agreement for participation. The 30 subjects were split into two groups; each group was tested by a different tester (20 subjects were tested by the author and the other 10 were tested by an MSc student). The reason for splitting the subjects over 2 testers was to look for any differences between the data obtained by the 2 testers. Each subject attended 2 sessions of around two hours each. At the beginning of Session 1, subjects underwent a hearing screening test (PTA) for the frequencies 0.25, 0.5, 1, 2, 4 and 8 kHz at 20 dB HL to ensure they had NH. Additionally, they filled out an otologic questionnaire to ensure they had no history/symptoms related to HL including recent ear disease, ear surgery, tinnitus, recent exposure to loud noise and family history of HL.

This study, which took place at the Institute of Sound and Vibration Research, University of Southampton, consisted of a pilot study followed by two main sessions. The same procedure was used in all sessions (the pilot study, Session 1 and Session 2). Subjects were seated in a soundproof room with a laptop placed on a table in front of them. An interface was built in the MATLAB software where all words were presented in pictures (Figure 4.5).

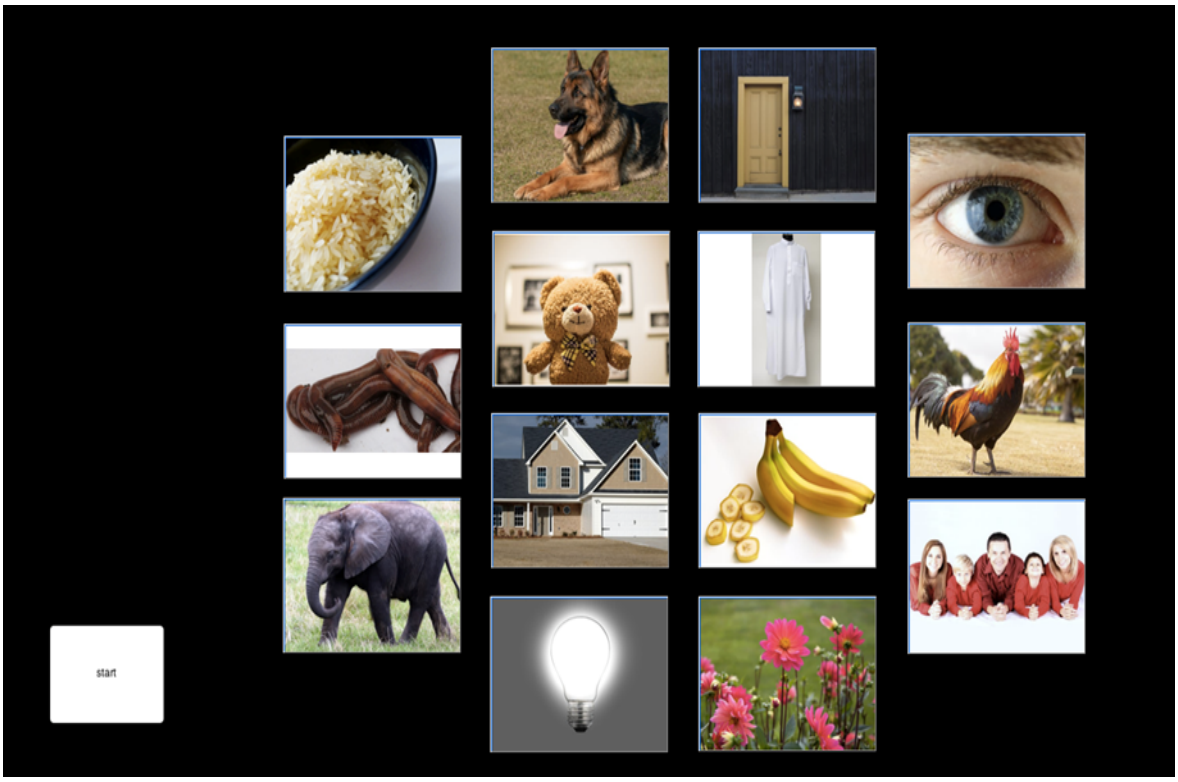


Figure 4.5 The MATLAB interface of the PAAST

For the testing, each word was concatenated to the carrier “Where is the…?” with a 60 ms silence gap between the word and the carrier phrase. A 500 ms silence gap was added prior to the carrier phrase and another 500 ms at the end of the sentence to avoid forward masking by allowing the noise to be ramped to the maximum presentation level before the speech starts. According to Moore (2003), forward masking decays to zero after at least 200 ms regardless of the initial amount of the forward masking. The onset and offset ramps of noise were 50 ms long. The silence gap allows the background noise to start and rise up to its steady-state level (60 dB A) prior to the beginning of the sentence and not to fall until the end of the sentence. A random segment/waveform of the generated stationary speech-spectrum noise that matches the length of the sentence was randomly selected and displayed in the background of each sentence. This was done to avoid using what is called a “frozen noise”, which would lead to better performance. Figure 4.6 illustrates a typical sentence of the PAAST.

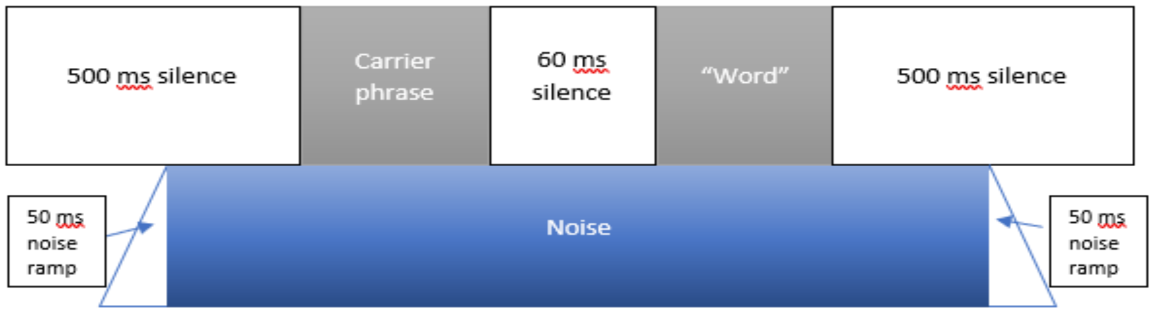


Figure 4.6 A typical sentence in the PAAST

Using the method of constant stimuli, the PAAST words were played in stationary speech-spectrum noise at different SNRs from the MATLAB via a sound card through Sennheiser headphones (the equipment was the same equipment used for the word recording, Section 4.3.2). Subjects were instructed to repeat each word once they heard it. When they were uncertain about the word played, they were encouraged to guess. The responses of the subjects were recorded automatically by the MATLAB software to be then used to fit the PFs of the words. Calibration was performed prior to starting the experiment in an artificial ear type 4153 using a flat plate coupler.

In Session 1 the words before adjustments were presented. The data obtained from the 30 subjects were then used to fit the PF of each word. The PFs were assessed in terms of the locations and slopes of the words and then the locations of deviated words from the mean locations were adjusted. Following the adjustments, the subjects were tested in Session 2 using the same procedure, but this time the words after adjustment were presented. After Session 2, the PFs of the words were fitted to assess the success of the adjustments in equalising the intelligibility of the words. The pilot study was conducted to find the best SNRs that would allow for responses from chance level to almost 100% correct responses. It was noted in Section 4.4.1 that careful selections of the SNRs would enhance the steepness of the slopes of the words and thus reduce threshold estimation errors. The following section outlines the pilot study.

4.4.2.1 Pilot study

Four NH Arabic-speaking adults were tested. The words were played in stationary speech-spectrum noise at the SNRs from −2 to −18 dB SNR in 2 dB steps resulting in 9 SNRs. Each word was repeated 5 times per SNR, leading to a total of 630 presentations of each subject. The presentation of the words was randomised by the MATLAB software and the SNRs were randomised by the tester. The software then recorded the responses for each subject.

By fitting the PF for each word, it was found that all words scored 100% at -2 dB SNR and zero at −18 dB SNR except for two words. The first is the word "sun" (/ʃʌms/), which scored 100% up to −16 dB SNR for all the four subjects. Apparently, the word "sun" (/ʃʌms/) was identified easily regardless of the stimulus level. This could be due to the presence of the two fricatives /ʃ/ and /s/ at the beginning and the end of the word. These two fricatives, /ʃ/ and /s/, are known to have high power (Gelfand, 2010), which would make them easily identifiable at low SNRs. The second is the word "fire" (/nɑr/), which showed a low percentage of correct responses even at high SNRs, indicating its difficulty to be identified in comparison to the rest of the words. These results indicated a large difference in the performance levels between these two words and the others. Therefore, it was decided to replace these words. By looking back at the list of words that was found to be familiar to the children, it was found that the word /wʌrd/, which means “flower”, has the same vowel as in the word “sun” (/ʃʌms/) and is 90% familiar to the children, so it was used to replace the word “sun” (/ʃʌms/). On the other hand, it was decided to exclude the pair /nɑɽ/ and /fɑɽ/ and replace it with another pair of words. It was thought that since some changes were being made, it was worth including a second pair of words with a short vowel to the list. The new pair of words is /dob/ which means “bear” and /roz/ which means “rice”, which share the short vowel /o/. As a result, the final word-list included five long vowels and 2 short vowels (/a, I, e, ɔ, u, ʌ, u/) as shown in Table 4.7.

Table 4.7 The final word-list of the PAAST

Pair	Arabic word	Phonetic transcription	English meaning
1	باب	/bab/	door
	ناس	/nas/	people
2	فيل	/fil/	elephant
	ديك	/dik/	hen
3	بيت	/bet/	house
	عين	/ʕen/	eye
4	ثوب	/θɔb/	Arabic men’s dress
	موز	/mɔz/	banana
5	دب	/dob/	bear
	رز	/roz/	rice
6	ورد	/wʌrd/	flower
	كلب	/kʌlb/	dog
7	دود	/dud/	worms
	نور	/nuɽ/	light

The newly selected words were recorded by the same speaker, using the same equipment (see Section 4.3.2). Afterwards, re-piloting was done where the same four subjects were tested using the new words. Since the piloting was repeated, it was decided to use different SNRs (−3, −5, −7, −9, −11, −13, −15 and −17 dB SNR) to allow for comparing the performance of the subjects over a wider range of SNRs (9 SNRs in the first pilot study and 8 SNRs in the second pilot study). This allowed for selecting the best SNRs that led to the monotonicity of the words and steeper slopes to be used in the main experiment. The number of trials was also increased to 10 trials per word at each SNR, resulting in a total of 1,120 presentations of each subject.

The findings of the re-piloting showed that all words scored zero at -18 dB SNR and 100% at -3 dB SNR. Based on this, it was decided to use the SNRs from -3 to -18 dB SNR in 3-dB steps (−3, −6, −9, −12, −15 and −18 dB SNR) for the purpose of the main experiment. Without any intention of doing so, the selected SNRs were the same SNRs used by Summerfield *et al.* (1994), who equalised the intelligibility of the IHR-McCormick Automated Toy Discrimination Test.

4.4.2.2 Study 3: methods

The method of constant stimuli, which ensures a presentation of a fixed number of each stimulus at each SNR, was used to present the words to the subjects. Each target word was played 10 times at each SNR (−3, −6, −9, −12, −15 and −18 dB). In total, 840 sentences were played to each subject (14 words x 10 trials x 6 SNRs). To vary the SNRs, both the noise and the speech stimuli levels varied equally but in opposite directions. For example, at −12 dB SNR, the noise level was increased by 6 dB, and the speech stimulus decreased by 6 dB. At an SNR of 0 dB, both the speech stimuli and the noise were set to 60 dB A. Subjects were asked to repeat the word they heard even if they had to guess. The MATLAB software recorded the responses given by each subject.

After completing Session 1 for all subjects, a logistic function model was used to fit the PFs of each word across all subjects through the MATLAB software (details of the mathematical equation used are described below). After fitting the PFs of the words, the locations, slopes and probability of deviance (pDev) values of each word were assessed. Assessing pDevs allows for checking the suitability of using the selected function to fit the PFs for a given data. The closer the pDev value to 1, the better the fit. PDev values that are less than 0.05 is an indication that the fit is unacceptably poor (Kingdom and Prins, 2010).

The intelligibility of the speech stimuli was equalised based on the locations, which is the steepest point on the x-axis, for the reasons discussed in Section 4.4.1. Following the adjustments, all subjects attended Session 2 where they were tested again using the same procedure used in Session

1, but this time the words were played after adjustments. The aim of conducting Session 2 was to ensure that the locations of the words after the adjustments showed homogeneity. Ethics approval was provided by the ethics committees at the University of Southampton, UK (identification number 26039).

Data analysis

Data were analysed using the MATLAB software. The MATLAB Palamedes toolbox (Kingdom and Prins, 2010) was used to fit the PFs of the words and to generate the location, slope and pDev for each word. The logistic function was used to fit the PFs of the words using the maximum likelihood criterion. The use of the logistic function is recommended for speech intelligibility experiments as it shows a reasonable sigmoidal model (Wichmann and Hill, 2001). Additionally, it showed a good fit when used for the same purpose in several previous studies (Summerfield *et al.*, 1994; Smits *et al.*, 2004; Potgieter *et al.*, 2016; Semeraro *et al.*, 2017; Yuen *et al.*, 2019). The logistic function that was used in the current study is given by Kingdom and Prins (2010) as:

$$f(SNR) = \frac{1}{1 + \exp\left(-\beta(SNR - x)\right)}$$

where parameter β corresponds to the slope and parameter x corresponds to the location. The function PAL_PFML_Fit was used to find the best fitting parameters. For that purpose, the guess rate was fixed to 1/14 and the lapse rate was fixed to zero (the appropriateness of this decision is discussed below in Section 4.4.3.1). The goodness of fit was estimated using the function PAL_PFML_GoodnessOfFit as described in Wichmann and Hill (2001).

Other statistical analyses were carried out using IBM SPSS statistics software version 24. Summary statistics were described using means and standard deviations (SD). To compare Data 1 (obtained from 20 subjects) and Data 2 (obtained from 10 subjects) in terms of the locations and slopes, the independent sample t-test was used. Additionally, to compare between the locations of the Combined Data (30 subjects) in Session 1 and Session 2, the repeated measures t-test was used. The selection of the test depended on the distribution of the data. Because no evidence was there to reject the null hypothesis (normal distribution of data), the parametric statistical tests were used.

4.4.3 Results

After selecting the words for the PAAST, it was necessary to equalise the intelligibility of the words to prepare it to be utilised in an adaptive procedure. For that purpose, two groups of subjects, who share the same inclusion criteria, were tested by two researchers using the same equipment and procedure. Based on this, two sets of data were available; Data 1 (given by 20 subjects) and Data 2 (given by 10 subjects). The raw data that represent the location of each word given by all subjects were treated by the MATLAB to fit the best-fit PF for each word. The flowchart below (Figure 4.7) shows how the results of the current study will be presented.

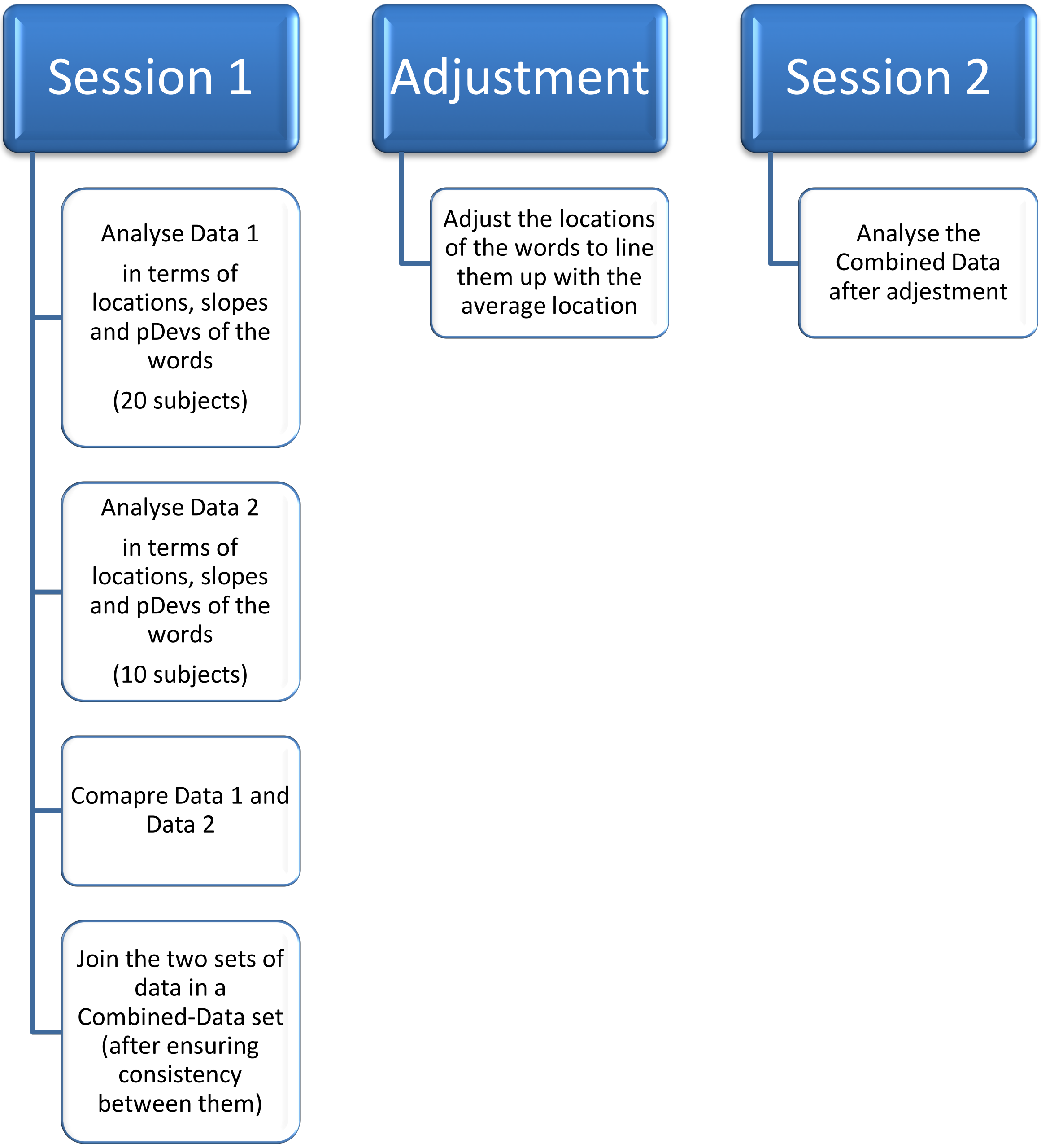


Figure 4.7 Flowchart of the results of Study 3

4.4.3.1 Session 1

Data 1

The data collected from 20 subjects were treated by the MATLAB software where the PF of each word was fitted. As it appears in Figure 4.7, the locations of the words are spread over a relatively large area on the x-axis (mean= −11.0 dB SNR, 95% CI of mean: −12.1 to −9.8 dB SNR, SD= 2.0 dB) with the word "people" showing the minimum location (−14.7 dB SNR) and the word "worms" showing the maximum location (−7.4 dB SNR). Additionally, it is obvious in Figure 4.8 that the word "rice" has the shallowest slope (β 0.4), and the word "people" has the steepest slope (β 1.3). The generated pDevs showed that all words have pDevs of >0.05, which is acceptable according to Kingdom and Prins (2010), with 11/14 words having a pDev ≥0.8. This showed that the usage of the logistic function was appropriate for fitting the PFs of the words (see Appendix F for the locations, slopes and pDevs of all the 14 words obtained from Data 1).

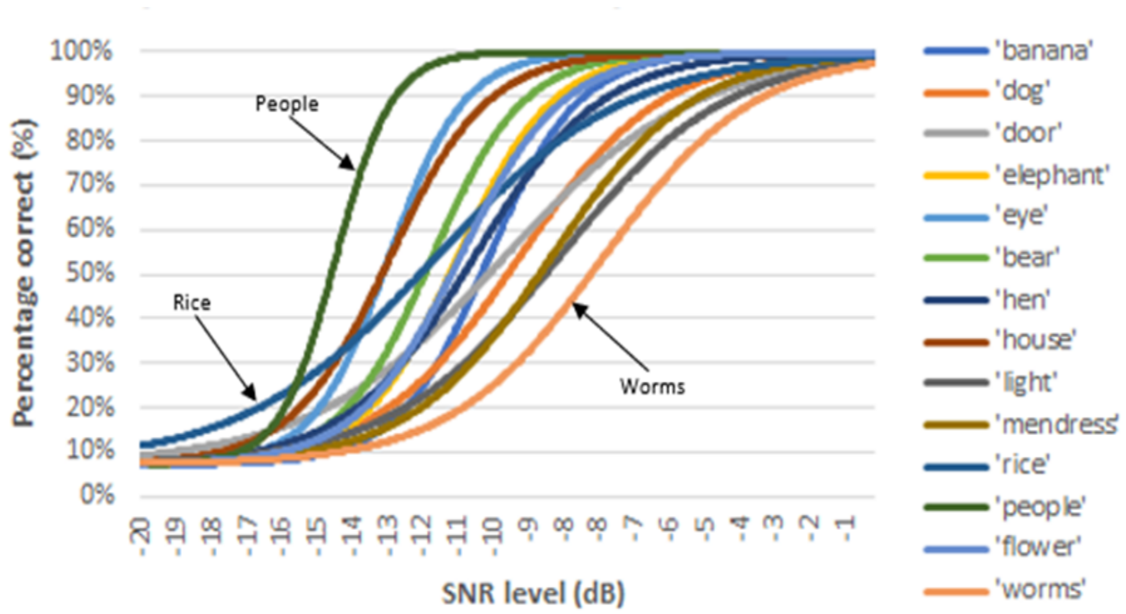


Figure 4.8 PFs of the words of the PAAST (Data 1)

Data 2

The data collected from another 10 subjects were used to fit the PFs of the words. It can be seen in Figure 4.9 that the spread of the locations of the words over the x-axis looks similar to the spread of the locations obtained from Data 1 (mean= −11.3 dB SNR, 95% CI of mean: −12.4 to −10.2 dB SNR, SD= 1.9 dB). Similar to Data 1, the word "people" showed the minimum location (−14.5 dB SNR) and the word "worms" showed the maximum location (−8.4 dB SNR). Additionally, the same word, "rice", showed the shallowest slope (β 0.3) whereas both "eye" and "house" showed the steeper slopes (β 1.3). The pDev values of all words were acceptable (see Appendix F for the locations, slopes and pDevs of all the 14 words obtained from Data 2).

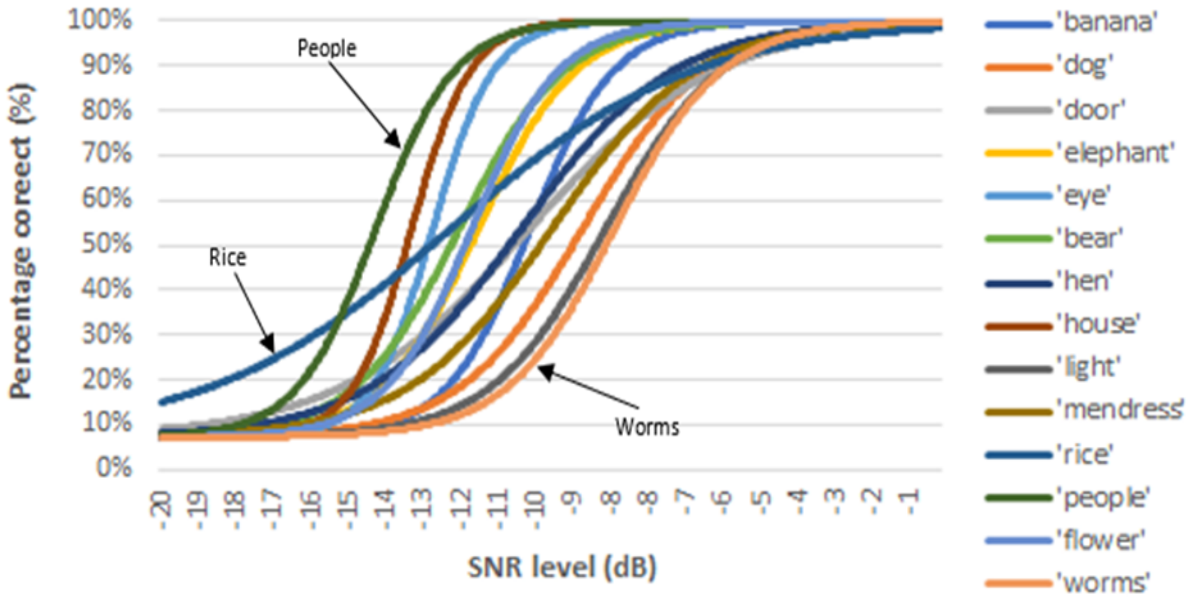


Figure 4.9 PFs of the words of the PAAST (Data 2)

Comparing Data 1 and Data 2

To assess the appropriateness of combining Data 1 and Data 2 and treat them as one group of data with a larger sample size (30 subjects), the two sets of data were compared statistically to each other in terms of the locations and the slopes of the words.

The mean locations of all words were compared between Data 1 and Data 2 using an independent sample t-test assuming equal variance since the data were at least approximately normally distributed (Shapiro-Wilk test: p= 0.9 for the two sets of data). No statistically significant difference was found between the mean locations of the two sets of data (t_{26}= 0.4, p= 0.7) with a mean difference of 0.3 dB (95% CI: −1.2 to +1.8 dB).

The mean slopes of all words were compared between Data 1 and Data 2 using an independent sample t-test assuming equal variance since the data were at least approximately normally distributed (Shapiro-Wilk test: p= 0.08 for Data 1 and p= 0.5 for Data 2). No statistically significant difference was found between the mean slopes of the two sets of data (t_{26}= −0.8, p= 0.4) with a mean difference of β −0.09 (95% CI: −0.3 to +0.1).

Since the results showed no statistically significant differences between Data 1 and Data 2 in terms of the locations and slopes of the words, the two sets of data were combined as one group of data, called the Combined Data. From this point, only the results of the Combined Data will be reported.

Combined Data

The PFs of the data collected from a total of 30 subjects were fitted (Figure 4.10). The pDevs were checked for all words to ensure the appropriateness of using the logistic function to fit the PFs of the words. Twelve out of 14 words have pDev values ≥0.7 (the lowest was 0.5). Thus, it showed that the usage of the logistic function was appropriate. The mean location of the words was −11.1 dB SNR (95% CI of mean: −12.2 to −9.9 dB SNR, SD= 2.0 dB). The minimum location was −14.6 dB SNR, and the maximum location was −7.6 dB SNR. The mean slope was β 0.7 (95% CI of mean: 0.5 to 0.8 β, SD= β 0.3). (see Appendix F for the locations, slopes and pDevs of all the 14 words obtained from the Combined Data).

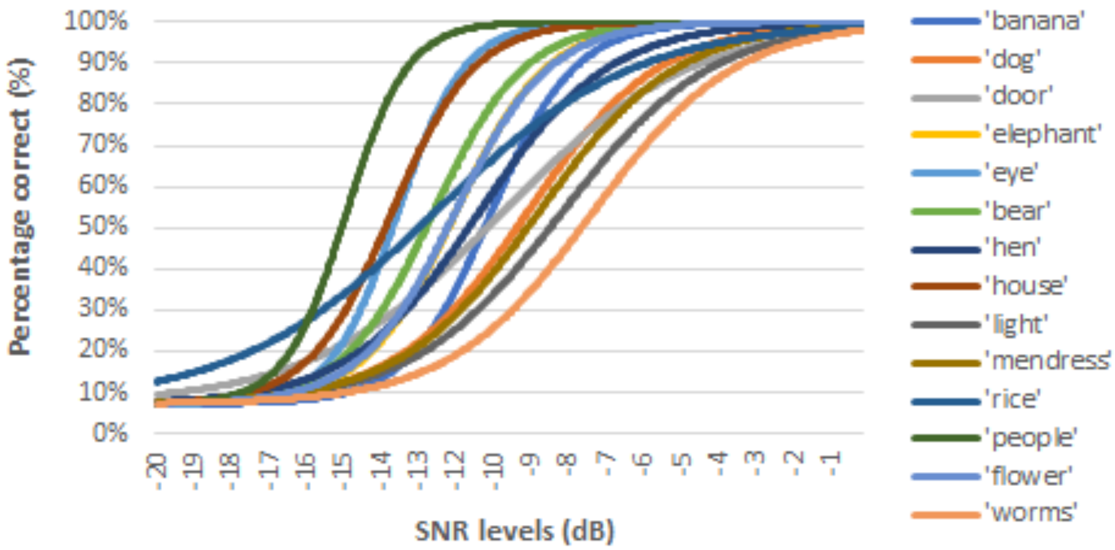


Figure 4.10 PFs of the words of the PAAST (Combined Data)

Some researchers recommended not fixing the lapse rate to zero but rather to a small number such as 0.01 because even when the speech stimulus is expected to be intelligible, subjects might not respond correctly 100% of the time due to some factors such as loss of attention, sneeze etc. Therefore, researchers advocate fixing the lapse rate to a small number to avoid bias (Klien, 2001; Wichmann and Hill, 2001; Kingdom and Prins, 2010). However, in this study, the lapse rate was fixed to zero for two reasons. First, the obtained scores reached 100% at the higher SNRs. Second, the locations and the slopes with the lapse rates fixed to 0.01 and zero were calculated and compared, and almost no difference between the values was found (the differences were within ±0.2 (see Appendix F). Thus, for this specific data, it was appropriate to fix the lapse rate to zero.

4.4.3.2 Adjustment

After completing Session 1 and assessing the locations and slopes of all words, the locations of the deviated words were adjusted to line them up with the average location of the words. In other words, the amplitude of the words that were easier to identify was decreased and the amplitude of the words that were harder to identify was increased. The required amount of adjustment was rounded to the nearest half decibel (see Appendix F). The maximum amount of adjustment was −3.5 dB for the word "worms" and +3.5 dB for the word "people". The adjusted words were then saved in new .wav files to be used in Session 2.

4.4.3.3 Session 2

Data were collected from the same 30 subjects but this time using the adjusted words. The PFs were fitted and assessed. As it appears in Figure 4.11, the gap between the locations of the words on the x-axis is narrowed in session 2. The maximum and minimum locations after adjustments were −10.5 dB SNR and −12.2 dB SNR respectively (mean= −11.4 dB SNR, 95% CI of mean: −11.6 to −11.1 dB, SD= 0.4 dB) (see Appendix F for the locations, slopes and pDevs of all words in Session 2).

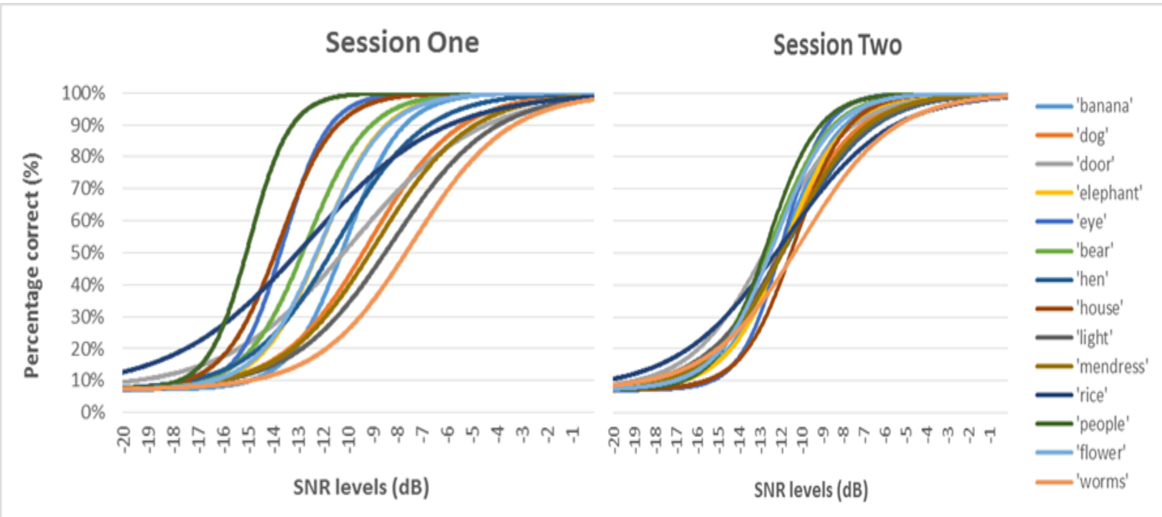


Figure 4.11 PFs of the words of the PAAST (Combined Data) Session 1 and Session 2

To assess the success of the adjustments in minimising the variations in the locations of the words, a comparison between the locations of the words in the two sessions was carried out by looking at the differences in the variances between the locations of the words in session 1 and Session 2. The variance in Session 1 was 4.0 dB, whereas the variance in Session 2 was 0.2 dB. The F test of variance indicated a statistically significant difference between the variance of the locations in the two sessions, where the locations in Session 1 vary more than the locations in Session 2 (calculated F=

20, critical value F=1.84 at α=0.05). This can be seen in the SD of the locations in the two sessions where the SD in Session 2 (0.4 dB) is much smaller than it was in Session 1 (2.0 dB), which indicates that the variations between the locations of the words were reduced.

To ensure that the mean location of the sessions was not affected by the adjustments, the mean locations of the words in Session 1 and Session 2 were compared using a repeated measures t-test. The parametric statistical test was used because the locations of the words in both sessions were at least approximately normally distributed (Shapiro-Wilk test: p= 0.9 for Session 1 and p= 0.5 for Session 2). The results revealed no statistically significant difference between them (t_{13}= 0.7, p= 0.5).

The adjustments that were made to the locations of the words were not expected to influence the slopes. Table 4.8 provides the mean and SD of the slopes in the two sessions.

Table 4.8 The mean and SD of the slopes in Session 1 and Session 2

	Mean (95% CI)	**SD**
Slope (β) (Session 1)	0.7 (0.5-0.8)	0.3
Slope (β) (Session 2)	0.6 (0.5-0.7)	0.2

4.4.4 Discussion

The current study aimed to equalise the intelligibility of the PAAST's speech stimuli in order to prepare the words to be utilised in an adaptive procedure to be used for threshold estimation. To achieve that, it was necessary to ensure that the words are homogenous, show a monotonic relationship and show steep slopes. For that purpose, the intelligibility of the PAAST words was assessed in stationary speech-spectrum noise. Two main features of the PFs of the words were measured. The first was the locations of the words, which is the SNRs at which the steepest point of the PF curve is found. The second was the slopes of the PF curves. As mentioned previously, the steeper the PF curve, the higher the precision of the SIN test in threshold estimation (Kingdom and Prins, 2010).

For the purposes of the study, data were collected by two researchers using the same method from two different samples that shared the same inclusion criteria. Data 1 included 20 subjects, whereas Data 2 included 10 subjects. All subjects were tested over two sessions. Following Session 1, a statistical comparison between the results of the two data (Data 1 and Data 2) was carried out. No statistical difference was detected between the two sets of data either in terms of the locations nor in terms of the slopes. Based on that, the two sets of data were combined and used as one group of data with a larger sample size (30 subjects).

The characteristics of the PFs of the words that were obtained from Session 1 were measured. All words showed acceptable pDev values (>0.05). The pDev values in the current study, which ranged from 0.75 to 0.99, are comparable to the pDevs reported by Semeraro (2015), who equalised the intelligibility of the words of the British English coordinate response measure (CRM) SIN test, where the pDevs ranged from 0.61 to 1. No other study was found to report the pDev values of their data. The high pDevs found in this study indicated that the logistic function was appropriate to fit the PFs of the words of the PAAST.

The results of Session 1 showed variations in the locations of the words, indicating that an adjustment was necessary to be made to minimise the variations. Most studies that attempted to equalise the intelligibility of speech stimuli manipulated a threshold point (SRTn) on the x-axis of the PF curve to equalise the intelligibility of the speech stimuli (Plomp and Mimpen, 1979; Summerfield *et al.*, 1994; Smits *et al.*, 2004; Ozimek *et al.*, 2009; Vaez *et al.*, 2014; Potgieter *et al.*, 2016; Semeraro *et al.*, 2017; Yuen *et al.*, 2019). No study in literature was found to manipulate the location to equalise the intelligibility of the speech stimuli. However, in the current study, the intelligibility of the words was equalised based on the locations of the words for the reasons discussed previously in Section 4.4.1. The maximum amount of adjustment was ±3.5 dB SNR. The two words that required the maximum adjustments were "people" (/nas/) and "worms" (/dod/). The phonetic structures of the two words in terms of the included sounds/phonemes were believed to be the reason for making them easy/difficult to identify. The word "people" (/nas/), which showed the minimum location (−14.6 dB SNR), includes a high-frequency sound (/s/), which would make it easily identifiable by NH listeners. On the other hand, the word "worms" (/dod/), which showed the maximum location (−7.6 dB SNR), has a voiceless sound, /d/, at the beginning and the end of the word. This makes it not surprising that the word becomes difficult to identify, particularly in the presence of background noise. However, it was decided not to exclude the two words from the test speech stimuli but instead adjust their locations because they did not deviate largely from the average location.

The amount of required adjustment is comparable to the amount of adjustment done by other researchers such as Lovett *et al.* (2013) who equalised the intelligibility of the words of the IHR-MTT (−2.5 dB SNR) and less than the maximum amount of adjustments done by other researchers who used words as their speech material. For example, the maximum amount of adjustment reported by Semeraro *et al.* (2017), Smits *et al.* (2004) and Yuen *et al.* (2019) were +4.08 dB SNR, +6 dB SNR and −5.8 dB SNR respectively.

Homogeneity is the first assumption that needs to be met before utilising the speech stimuli in an adaptive procedure. The key finding of this study that indicates homogeneity of the PAAST words is the reduction in the SD of the locations of the words from 2.0 dB in Session 1 to 0.4 dB in Session 2 (this can be seen in Figure 4.11 as the gap between the words is narrowed in Session 2). This finding was supported statistically by the finding of the F test, which revealed statistically significant differences between the variance of the locations in the two sessions. The findings suggest that the adjustments to the locations obtained from Session 1 were successful in enhancing the homogeneity between the words in Session 2.

Monotonicity of the speech stimuli is the second assumption that needs to be met before utilising the speech stimuli in an adaptive procedure. A monotonic relationship can be seen clearly in Figure 4.11 with the increase in SNR results in an increase in performance levels and vice versa.

The steepness of the slopes is the third assumption that needs to be met before utilising the speech stimuli in an adaptive procedure: the steeper the slope, the more precise the threshold estimation. No information was found in the literature to report what values of slopes could be considered steep enough to estimate thresholds precisely. A systematic review that investigated the difference in slope values between different studies, which focused on fitting PFs for speech stimuli, reported a large variation in the accepted slopes between different studies; 1%/dB to 44%/dB (mean= 7.5%/dB). This variation was drawn back to several factors, including the type of speech stimuli, the type of noise and the number of maskers (MacPherson and Akeroyd, 2014). Therefore, a comparison between the slopes of different studies should be done with caution. Besides, the use of different equations when fitting the PF of the speech stimuli would lead to different slope calculations. Therefore, the slope values found in the current study were compared to the slopes reported by two studies which used almost the same equation used in this study.

In the current study, the slope was calculated based on the following equation (Kingdom and Prins, 2010):

$$f(SNR) = \frac{1}{1 + \exp\left(-\beta(SNR - x)\right)}$$

Semeraro (2015) used the same equation to calculate the slope; therefore, it was proper to run a comparison between the findings of the two studies. All the words in the current study showed fairly steep slopes (>β 0.5) except for two words, "worms" and "rice", where the slope was equal to β 0.4 for both words. Semeraro (2015) accepted a comparable slope (β 0.37) and reported reliable findings when the CRM was utilised in an adaptive procedure, indicating that a slope of around β 0.4 has no negative effect on threshold estimation.

Smits *et al.* (2004), who equalised the intelligibility of the Dutch triple-digit test in noise, used the following equation to fit the PFs of the speech stimuli and calculate its parameters:

$$f(SNR) = \frac{1}{1 + \exp\left(-4\beta(SNR - x)\right)}$$

They rejected stimuli that showed a slope of less than 9%/dB. By calculating the slopes of the words of the PAAST using the equation they used, it was found that the minimum slope in this study was 10%/dB and the maximum was 22.5%/dB; hence it is comparable to the slopes accepted in Smits *et al.* (2004).

In sum, it seems that slopes of the words of the PAAST are considered steep enough in comparison to the slopes of the speech stimuli of other SIN tests and are appropriate to be used to estimate SRTs. Thus, the third objective of the current study was achieved.

4.4.5 Conclusions

The current study aimed to equalise the intelligibility of the speech stimuli of the PAAST by ensuring that the PFs of the words show homogeneity, monotonicity and have steep slopes. To achieve this aim, the locations of the words were manipulated to match the average location of the words. Based on the findings, it was concluded that:

- The adjustments were successful and resulted in homogenous speech stimuli
- The steepness of the slopes of the words were acceptable when compared to previous studies.

To conclude, the speech stimuli of the PAAST are now almost equally intelligible, and the test is ready to be utilised in an adaptive procedure.

Chapter 5 Study 4: Assessing the reliability of the PAAST when utilised in an adaptive procedure (normal-hearing adults)

5.1 Overview of the chapter

In the previous chapter, the intelligibility of the speech stimuli of the Paediatric Arabic Auditory Speech Test (PAAST) was equalised. At this stage, the speech stimuli of the PAAST was utilised in an adaptive procedure, which is a common method of measuring speech reception thresholds (SRTs), and the PAAST was implemented in an iPad application to ease its accessibility and to appeal to children, as discussed in Chapter 2, Section 2.4.3. The current chapter explores the test-retest reliability of the PAAST in estimating SRTs. The reliability of the PAAST implemented in the iPad application was assessed against its reliability when performed through a MATLAB software using tried-and-tested code (considering the MATLAB version as an ideal). The SRTs obtained from normal-hearing (NH) Arabic-speaking adults were measured for that purpose (Study 4).

5.2 Introduction

One common method of estimating SRTs is an adaptive procedure. Adaptive procedures, as discussed in Chapter 2, Section 2.3.3.1, increase the efficiency of a test procedure and reduce the effort of the tester and the subject by preventing the collection of data that would not be informative (Kingdom and Prins, 2010). Therefore, the speech materials of the PAAST were utilised in an adaptive procedure in order to measure the SRT of Arabic-speaking individuals. The characteristics of the adaptive procedure that were used in the PAAST are discussed below in Section 5.3.1.

As mentioned in Chapter 2, Section 2.4.3, the aim was to implement the PAAST in an iPad application in order to ease the accessibility to the test, thus widening its usage in the KSA and around other Arab countries without the need for special equipment. The implementation of the PAAST in an iPad application does not only ease the accessibility of audiologists to the test but also eases its usage by non-specialists for screening purposes, particularly as the PAAST was designed to meet the criteria of screening tools, which were noted in Chapter 4, Section 4.2. Therefore, at this stage, the PAAST was implemented in an iPad application.

It is essential to assess the reliability of a newly developed test in order to ensure that it produces consistent results over repeated measurements (Kumar, 2011). Ideally, each test is expected to lead to the same score when performed by the same subject under the same conditions on two different occasions, indicating no measurement error, which is the difference between the measured value and the true value (Bell, 2001). However, this is difficult to find in real practice because of several factors such as the variability in human performance, the tester experience, the instrument used etc. (Kumar, 2011; Bland, 2015). Therefore, finding some variation in the obtained results is expected; however, the observed variation should be minimised as much as possible.

Variations in subjects' performance might result from systematic errors and/or random errors. Systematic errors, as indicated by the name, are errors that occur systematically, such as errors resulting from fatigue or learning effects (Bell, 2001). For instance, the scores obtained might become worse after a few trials because of the subject's tiredness. In contrast, the scores obtained might improve with trials because the subject became more familiar with the task. Systematic variations are assessed by measuring the difference between the trials across the sample. On the other hand, random errors refer to random changes in the scores obtained from a subject over repeated measures (Bell, 2001). Although it is difficult to control random errors, they can be assessed and calculated mathematically by measuring the within-subject standard deviation and the repeatability of the scores obtained from repeated measures (Bland, 2015). Since observing some variation between the scores obtained from repeated measures is expected, the test is considered to be reliable if it shows an acceptable measurement error for the effective practical measurement of that specific test.

At this stage, the reliability of the PAAST was unknown. Therefore, the aim of Study 4 was to assess the reliability of the PAAST in measuring SRTs in noise. However, it was preferred at this stage to assess the reliability of the PAAST implemented in an iPad application with NH adults and to compare it against its reliability when performed via the MATLAB software. This is because the MATLAB settings were considered to be ideal since the MATLAB code used was tried and tested in the lab in several previous experiments (e.g. Semeraro *et al.*, 2017).

The reasons behind testing adults, although the PAAST was developed for children, are given below. First, in order to compare the reliability of the PAAST implemented in the iPad application against its reliability when implemented in MATLAB software, each subject had to perform the test several times in each condition (MATLAB and iPad). This was expected to be difficult to tolerate for children. Additionally, it would be easier to retest or recruit adult subjects if amendments needed to be made at this stage.

Second, with the lack of Arabic automated SIN tests, the PAAST can be used clinically to estimate the SRT of Arabic-speaking adults (e.g. adults with severe/profound HL or adults with cochlear implants). Ensuring that the PAAST provides reliable results with adults before using it in clinics is essential. Besides, to be used in clinical settings, audiologists would need information about the performance of NH Arabic-speaking adults in SIN testing to be used as a reference. There is a gap in knowledge about the performance of NH Arabic-speaking adults in SIN tests as no SIN test is available for adults in Arabic. Therefore, providing such information was considered at least to be used as a baseline for further research work.

The duration required to complete the PAAST was also estimated in this study. It is important to estimate the test duration to ensure that it could be used for screening purposes, which require the test to be quick. It is also preferable in clinical practice to have tests that could be completed in short durations in order to account for the clinic's workload.

5.2.1 Aims

The aims of Study 4 were to:

1. provide preliminary data about the performance of Arabic-speaking adults in SIN testing
2. assess the test-retest reliability of the PAAST utilised in an adaptive procedure with NH Arabic-speaking adults in two conditions: the test implemented in an iPad application against the test implemented in MATLAB software
3. estimate the duration required by NH adults to complete the PAAST.

5.3 Methods

5.3.1 Adaptive procedure

The adaptive procedure that was used is the 1-up/2-down method, which was first developed by Wetherill and Levitt (1965). This method estimates the threshold corresponding to 70.7% correct responses by increasing the stimulus level following each incorrect response but not decreasing it until two correct sequential responses occur (Kingdom and Prins, 2010). Although this method requires a longer time than the 1-up/1-down method (as the number of trials required to reach the near-threshold level is more), it has some advantages. First, it is more precautious as it reduces the probability that a correct response is given by chance, and thus it prevents reducing the stimuli below the threshold after several correct responses occur by chance. Second, Brand and Kollmeier (2002) reported that the reliability of an adaptive procedure rapidly increases when the number of responses required to elicit a change in the stimulus level is more than 1. However, to avoid having many presentations at intensities that are way above the threshold, Wetherill and Levitt (1965) suggested starting with the 1-up/1-down method until the first reversal of direction is needed and then continuing the test using the 1-up/2-down method.

The initial presentation level of the PAAST, at which the speech is expected to be heard clearly, was 5-7 and 5 dB SNR for the MATLAB and iPad respectively. For the following trials, the SNR was varied adaptively. At positive SNRs, the speech level was fixed at 60 dB A and the noise level was varied whereas at negative SNRs, the noise level was fixed at 60 dB A and the speech level was varied. This procedure avoids ceiling and floor effects and thus avoids uncomfortably loud presentations. It works for individuals who need positive SNR to respond correctly and for individuals who can provide correct responses at negative SNRs.

Following two correct sequential responses, the SNR was decreased by 8 dB until the subject responds incorrectly where the SNR was increased by 4 dB. Afterwards, the adaptive procedure of 1-up/2-down is started and continued for 8 reversals with a 2-dB step size (Figure 5.1). The SRT was then estimated based on the average of the last 8 reversals. The selected step size allows for starting with a high SNR, reaching the threshold level rapidly and estimating the threshold as accurately as possible. The selected termination criteria (8 reversals at a step size of 2-dB SNR) allow for having a minimum of 30 presentations as recommended by Brand and Kollmeier (2002) and Wetherill and Levitt (1965), cited in Levitt (1971), who reported that at least 6 or 8 reversals with 30 presentations are required to obtain reliable SRT estimation. However, in cases where the individual listens to 60 presentations before completing 8 reversals at 2-dB SNR, the test is also terminated. This termination criterion was set to avoid running the test infinitely with inconsistent individuals.

In the iPad application, the upper cap was fixed to +20 dB SNR, and the lower cap was fixed to −20 dB SNR. So, if a child failed to respond correctly up to an SNR of +20 dB, the test would stop automatically; otherwise, it would act as a speech-in-quiet test. Similarly, it is not reasonable to go lower than −20 dB SNR because NH listeners are not expected to identify the words of the PAAST at lower levels (the results of Study 3 showed that NH adults scored 0% at −18 dB SNR).

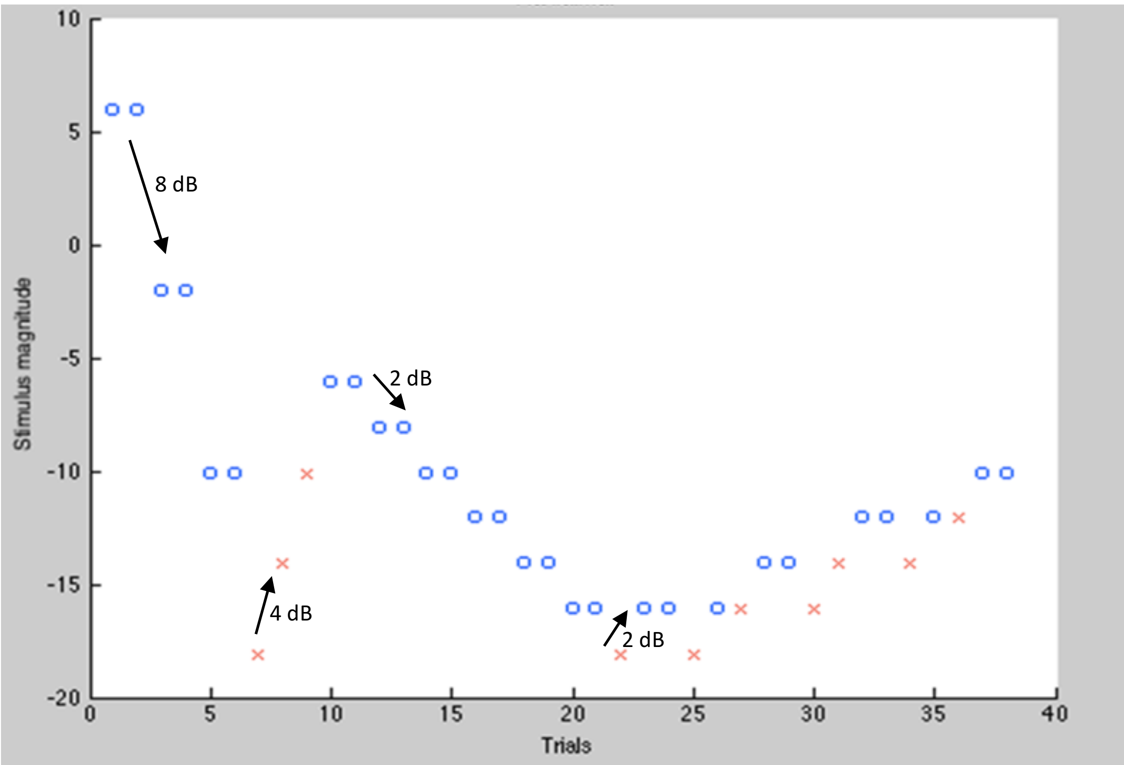


Figure 5.1 Graphical presentation of the adaptive procedure used in the PAAST (the circle signs represent correct responses, and the x signs represent incorrect responses)

5.3.2 Pilot study

A pilot study was carried out before proceeding to the main study. Five adults, who are native Arabic speakers and to their knowledge have NH, participated in the pilot study. They were screened for HL, and then their SRTs in noise were measured using the PAAST. Each subject attended one session where the PAAST was performed 3 times through the MATLAB software and 3 times through the iPad application. The test took place in a quiet room at the Institute of Sounds and Vibration at the University of Southampton.

The average SRT of the 3 rounds of the iPad was compared to the average SRT of the 3 rounds of the MATLAB across the 5 subjects. The mean SRT of both conditions were similar (SRT= −12.8 dB SNR, −12.9 dB SNR for MATLAB and iPad respectively) with only 0.1 dB difference between the SRTs obtained under the 2 conditions. The results indicated the readiness of the test to be used for the purposes of Study 4.

5.3.3 Study 4: methods

Arabic-speaking adults were recruited for the purposes of this study. The inclusion criteria were: (1) native Arabic speakers, (2) aged 18-50 years, and (3) have PTA thresholds ≤20 dB HL for the frequency range from 0.25-8 kHz. The exclusion criterion was: have HL, or any symptoms related to HL. Thirty adults (14 males and 16 females) aged 18-41 years old (mean= 27.4 years, SD= 6.7 years) were included. Written consent was obtained from the subjects prior to any testing. At the beginning of the session, subjects were asked to answer a short otological-health questionnaire in order to rule out any HL-related symptoms including recent ear disease, ear surgery, tinnitus, recent exposure to loud noise and family history of HL. Additionally, subjects were screened for HL using a GSI-61 audiometer and TDH headphones at the frequency range of 0.25-8 kHz. Testing took place in a quiet room at the Institute of Sounds and Vibration at the University of Southampton, where each subject attended one session that lasted about 35-45 minutes.

5.3.3.1 Sample size

The sample size necessary to assess the test-retest reliability was calculated based on a mean difference of 0.7 dB SNR and a SD of 1.3 dB, which is close to the findings of some previous studies (Summerfield *et al.*, 1994; Smits *et al.*, 2013; Sheikh Rashid *et al.*, 2017). It was found that 27 subjects are enough to detect a mean difference of 0.7 dB SNR with a power of 0.8 and an alpha of 0.05 (Howell, 1997). However, 30 subjects were recruited to consider for possible drop-offs.

5.3.3.2 Materials and calibration

For the purposes of the study, a MacBook Air laptop with the MATLAB software, an iPad with the PAAST, Sennheiser HAD 200 headphones, and a sound mixer (YAMAHA Mixing console with two microphone inputs, model MG06, serial number: BGXJ01035) were used. The laptop, the iPad and the headphones were connected to the sound mixer in order to process the output from the laptop and the iPad to the headphones.

Objective calibration was conducted before carrying out any data collection and repeated once weekly using a sound level meter (SLM), manufactured by Brüel & Kjær (type: 2260, serial number: 1853816), a sound level calibrator (94 dB, 1 kHz) manufactured by Brüel & Kjær (type: 4230, serial number: 861555) and an artificial ear (type: 4153, serial number: 198145).

5.3.3.3 Procedure

During the session, subjects were seated on a comfortable chair with the laptop/iPad placed on a table in front of them. Pictures representing the 14 words of the PAAST were displayed on the laptop/iPad screen in front of the subject (Figure 4.1 in Chapter 4 shows a screenshot of the MATLAB interface and Figure 5.2 below shows a screenshot of the iPad interfaces). Headphones were placed over the subject's ears, and the subject was instructed to listen carefully to the sentences played (in the form of "Where is the (word)?") and respond by clicking on/touching the picture that represents what was heard (Figure 5.3). Subjects were asked to guess if they were not sure what the correct answer was as the following word would not be played unless a response was recorded.

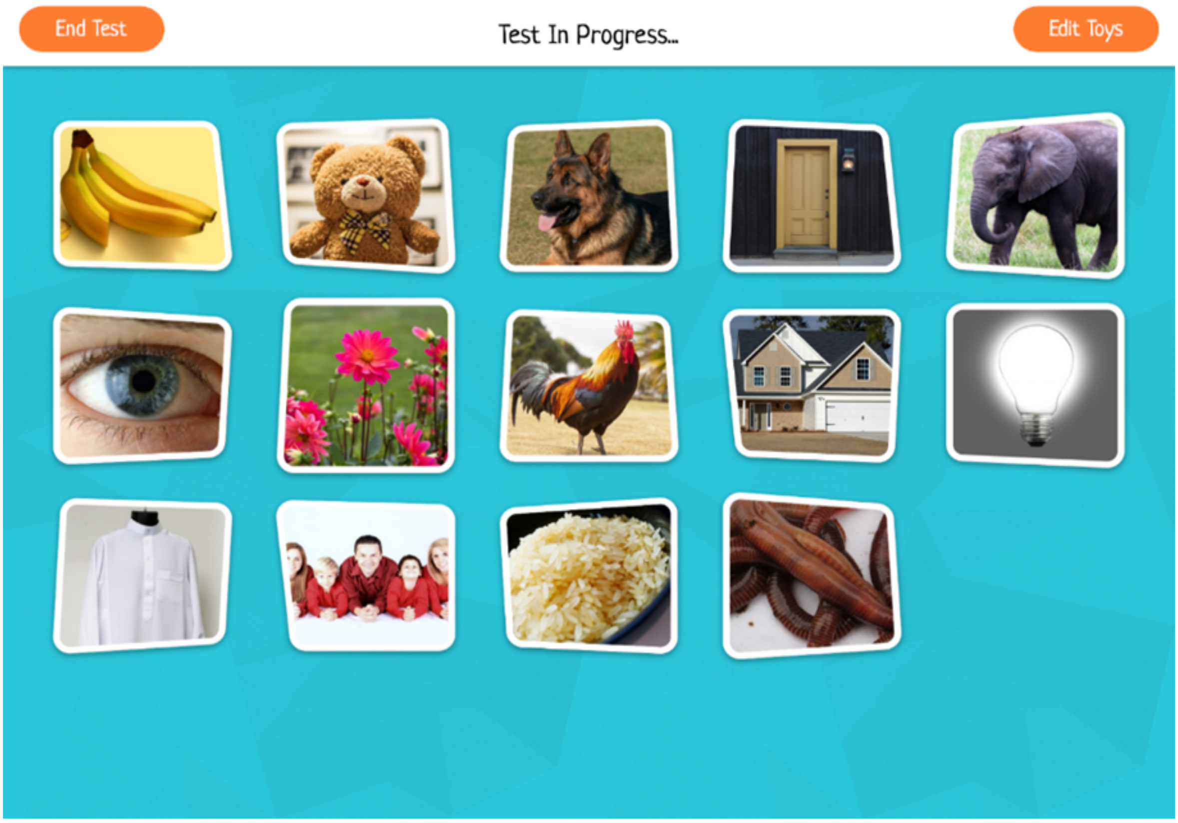


Figure 5.2 Screenshot of the iPad interface of the PAAST

Figure 5.3 Subject performing the PAAST (Study 4)

SRTs of the subjects were measured binaurally; 3 rounds with the MATLAB and 3 rounds with the iPad using the same adaptive procedure described earlier. All subjects completed 3 rounds with the MATLAB and then started with the iPad. The MATLAB software/iPad application randomised the presentation of the words, calculated the SRT based on the described adaptive procedure and recorded the results automatically. Subjects were offered breaks when they requested them.

Ethics approval was provided by the ethics committee at the University of Southampton in the UK (identification number 40367).

5.3.3.4 Reliability measurements

To assess the random error component of the test-retest reliability, the following three statistical measures were used as they are commonly used for this purpose (Plomp and Mimpen, 1979; Summerfield *et al.*, 1994; Lovett *et al.*, 2013). First, the **within-subject standard deviation (SDω)** of the scores obtained. Measuring the SDω, allows for assessing the ***variability*** in the results obtained from the same subject when the test is performed for several times. The smaller the SDω, the less variable the results, thus the more reliable the test.

The SDω is measured using the formula:

$$\sqrt{(\sum_i^n d^2)/2n},$$

where *d* is the difference between the scores obtained by the i^{th} subject and *n* is the number of subjects (Bland and Altman, 1996).

It is known that any single measurement obtained from a subject lies within ±1.96 SDω with a probability of ≥0.95. Thus, by measuring the SDω, one could estimate a subject's true value with a probability of ≥0.95 (Bland, 2015).

The second measurement is the ***repeatability*** of the scores, which is the minimum value that would tell that there is a statistically significant difference between any two measurements (Summerfield *et al.*, 1994; Lovett *et al.*, 2013). In other words, it is the smallest difference that can be reliably detected. This critical value is important in clinical practice. For instance, if a subject is tested in two different conditions (e.g. aided and unaided), and scored more than the repeatability/critical value, it can be concluded that the difference between the scores is statistically significant and did not occur by chance. So, if the SRT obtained from a subject when tested with a hearing aid was better than the SRT of the same subject when tested unaided by more than the amount of the repeatability, one could tell that this improvement is true and resulted from the use of the hearing aid. The repeatability of the scores is measured by multiplying the **SDω by √2 x 1.96** (Bland and Altman, 1996).

The third measurement is the ***correlation coefficient (r)*** between any two scores. The higher the correlation (large value of *r*), the more reliable the test. The intra-class correlation coefficient (ICC) is recommended for reliability assessment as it can be used when more than two trials are performed, and it does not consider the order of the trials (Bland, 2015).

5.4 Results

5.4.1 Mean SRT

The mean SRT across all rounds was −10.9 dB SNR (SD= 1.3 dB, 95% CI: −11.4 dB to −10.4 dB) for the MATLAB and −11.2 dB SNR (SD= 1.3 dB, 95% CI: −11.7 dB to −10.7 dB) for the iPad. The mean SRT for each round tested in both conditions (MATLAB and iPad) are shown in Table 5.1. Boxplots of the SRT across each round for both the MATLAB and iPad are shown in Figure 5.4.

Table 5.1 The mean SRT, SD and 95% CI of the mean provided through the MATLAB and iPad

		Mean SRT (dB SNR)	SD (dB)	95% CI of the mean (dB)
MATLAB	**Round-1 (R1)**	−10.4	1.7	−11.0 to −9.7
	Round-2 (R2)	−11.1	1.5	−11.6 to −10.5
	Round-3 (R3)	−11.4	1.7	−12.0 to −10.7
	Average of 3 rounds	−10.9	1.3	−11.4 to −10.4
iPad	**R1**	−10.9	2.0	−11.6 to −10.1
	R2	−11.4	1.4	−11.8 to −10.8
	R3	−11.5	1.3	−12.0 to −11.0
	Average of 3 rounds	−11.2	1.3	−11.7 to −10.7

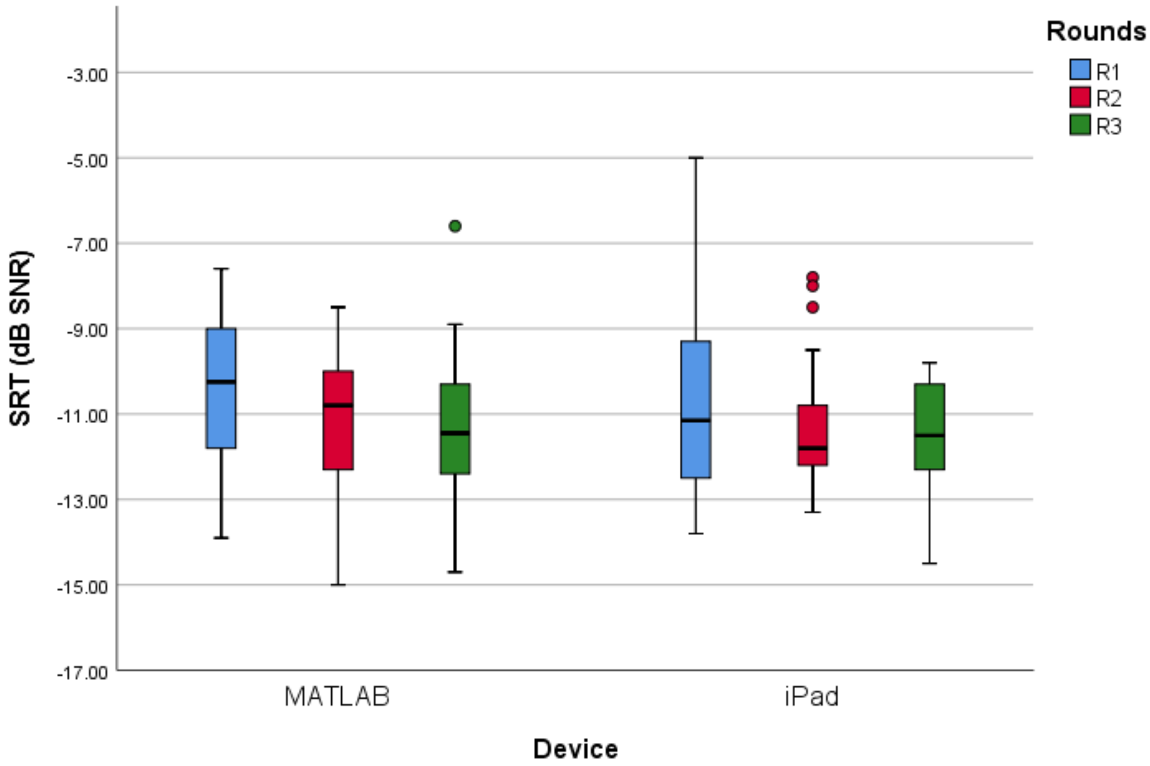


Figure 5.4 Boxplots representing the SRTs obtained through the MATLAB and iPad across each round

It can be seen in Figure 5.4 that the variability in the performances of the subjects in each round is relatively small except for iPad-R1 where more variation can be seen. By looking at the 3 rounds performed through the MATLAB, it can be seen that the performance of the subjects slightly improved in R3 in comparison to R1 and R2, which might indicate a learning effect. Four outliers were detected in the data, 1 in MATLAB-R3 and 3 in iPad-R2. When performing the statistical analysis, the outliers were not excluded in order to not bias the results.

5.4.2 Test-retest reliability

As mentioned previously, variations in the performance might result from systematic or random errors. In order to look for a systematic error in the performance of the subjects, the mean SRTs obtained in the 3 rounds of the 2 conditions (MATLAB and iPad) were compared using the repeated measures analysis of variance (ANOVA), considering the SRT as the dependent variable, the conditions and the rounds as the factors and the number of rounds as the levels (3 rounds for each condition). Data for the 3 rounds obtained through the MATLAB were at least approximately normally distributed (Shapiro-Wilk test: p= 0.2, p= 0.4 & p= 0.4 for R1, R2 and R3 respectively). When subjects were tested through the iPad, data obtained in R3 were at least approximately normally distributed (Shapiro-Wilk test: p= 0.1), but data obtained in R1 and R2 did not show normal distribution (Shapiro-Wilk test: p= 0.02 and p= 0.005 for R1 and R2 respectively). However, it is known that ANOVA is a robust statistical test (Field, 2013, p 444); thus it could still be used even with data that do not show statistical normal distribution. Besides, Levene's test of equality of error variances indicated that the homogeneity variance assumption was not violated and the error variance of the dependent variable (SRT) is equal across all rounds (p=0.1). Based on this, it was decided that it is appropriate to conduct the repeated measures ANOVA on the data.

Mauchly's test of sphericity indicated that the assumption of sphericity was met for the rounds (X^2 (2) = 2, $p \geq 0.1$) but violated for the interaction (X^2 (2) = 8.8, p= 0.01). However, it does not apply for the conditions as only 2 conditions are there in this study (MATLAB and iPad). Therefore, the two-way repeated-measures ANOVA with sphericity assumed was used to find the main effect of the rounds and the two-way repeated measures with Greenhouse-Geisser correction was used to find the main effects of the interaction on the SRT.

The main effect of the 2 conditions was not statistically significant, F (1, 58) = 0.9, $p \geq 0.1$ with a mean difference of 0.3 dB (95% CI: −0.3 to 0.9 dB). The mean SRT obtained through the iPad across the 3 rounds was 0.3 dB higher than the mean SRT obtained through the MATLAB. By looking at the 95% CI of the mean, a difference between the 2 conditions to less than 1 dB could be ruled out, to 95% certainty. The main effect of the rounds was statistically significant, F (2, 58) = 6.9, p= 0.002 in contrast to the main effect of the interaction between the conditions and the rounds, which was not statistically significant, F (2, 58) = 0.5, *p*=0.5. This indicated that the effect of the rounds does not depend on the condition and vice versa. Table 5.2 provides the mean differences between the 3 rounds across the 2 conditions.

The post hoc comparison test was carried out with a Bonferroni correction. The only statistically significant difference was found between R1 and R3 (*p*= 0.004) with a mean difference of 0.8 dB (95% CI: 0.2 to 1.3 dB). The SRT of R3 was better than that of R1 by 0.8 dB. Although the difference was statistically significant, a difference to just above 1 dB (1.3 dB) can be ruled out with 95% certainty. This difference could probably be due to a learning effect, which is not uncommon and is expected to occur with repeated measurements. The results indicated no statistical difference between the subjects' performance when tested using the MATLAB and iPad settings.

Table 5.2 The mean differences (and 95% CI of the mean difference) between the SRT of the three rounds across the two conditions (MATLAB and iPad)

Rounds	Mean difference (dB SNR)	95% CI (dB SNR)	*p*-value
R1 vs R2	0.5	−0.04 to 1.1	0.08
R1 vs R3	0.8	0.2 to 1.3	0.004
R2 vs R3	0.2	−0.2 to 0.7	0.6

To look for the random error, the variability, repeatability, and correlation coefficient between the scores were measured. The random error was assessed for the iPad setting only since it is the one that is going to be used in real practice. The **variability** of the SRT scores was assessed by measuring the SDω of the SRT scores obtained from the subjects. The measured SDω was equal to 1.6 dB. As mentioned previously, a single score from any random subject is expected to lie within ±1.96 SDω with a probability ≥0.95 (Bland, 2015). So, any one measurement with the PAAST will lie within ±3.1 dB of the individual true threshold.

The **repeatability** of the test was 4.4 dB. Thus, a difference of >4.4 dB between the scores obtained in two different conditions (e.g. aided and unaided) is considered a true significant difference. Another measurement that was performed to assess the test-retest reliability of the PAAST is the **correlation coefficient** of the scores between the different rounds. This was assessed using the ICC, which calculates an average correlation coefficient of the measures of multiple estimators. For this data, the estimators were the rounds. The ICC with the two-way mixed model was carried out to measure the absolute agreement. The use of the two-way mixed model indicates that all the available estimators are included in the test. The correlation coefficient between the SRT of the 3 rounds was 0.7 (95% CI: 0.7 to 0.9, $p<0.0001$). This suggests a strong positive correlation between the rounds (Mukaka, 2012), indicating that the PAAST provides reliable SRTs for adults.

5.4.3 Test duration

The mean duration required to complete the test was 2.1 minutes (SD= 0.3 minutes) for the MATLAB and 4.6 minutes (SD= 0.5 minutes) for the iPad. The two sets of data did not show a statistically normal distribution (Shapiro-Wilk test: $p= 0.04$ and $p= 0.01$ for the MATLAB and iPad, respectively). Therefore, the non-parametric statistical test, Mann-Whitney test, was carried out and determined a statistically significant difference between the duration needed to complete the test with the two conditions (U= 0, $p< 0.0001$).

5.5 Discussion

The current study aimed to provide data about the performance of NH Arabic-speaking adults with the PAAST, to assess the test-retest reliability of the PAAST implemented in an iPad application against its reliability when performed through MATLAB software, and to assess the duration of the PAAST. SRTs of NH adults were measured by the PAAST for the purposes of the study.

5.5.1 Mean SRT

The mean SRT obtained by NH adults did not show a statistical difference when measured through the MATLAB and the iPad. In other words, performing the PAAST through the iPad application leads to similar SRTs as if the test were administered through the MATLAB. This finding is reassuring as it indicates that the PAAST works as expected when administered through the iPad application. Because the ultimate interest in real practice is to administer the PAAST through the iPad application, the discussion of this section focuses on the results obtained from the iPad testing.

To the best of the author's knowledge, no published SIN test is available for Arabic-speaking adults. Therefore, the comparison was made with the findings of studies that developed adults' SIN tests in other languages. The mean SRT obtained from the iPad testing is −11.2 dB SNR. A comparable finding was reported by Potgieter *et al.* (2016) who validated the South African smartphone-based digit-in-noise test and found a mean binaural SRT of −11.7 dB SNR when Sennheiser headphones were used. In addition, Wilson and Weakley (2004), who measured the word recognition ability in multi-talker babble using digit-triplet stimuli, reported a comparable result (−12 dB signal-to-babble). However, most of the other previous studies tested their subjects monaurally whereas the subjects in Study 4 were tested binaurally, which makes the direct comparison with their findings inaccurate because it is known that binaural listening leads to better SRTs than monaural listening. For instance, Plomp and Mimpen (1979) reported a difference of 1.4 dB between the average SRT of the monaural and the binaural testing (binaural provides better SRTs than monaural). In line with that, Wilson (2003), who developed the word recognition in noise test, reported a difference of 1.7 dB between the monaural and binaural testing.

Because in clinical practice it is preferable to perform monaural testing as well as binaural testing, it was decided to estimate the mean monaural SRT for NH adults when tested by the PAAST. For that reason, a study was conducted in Palestine (following Study 4) by collaborator researchers using the PAAST iPad application. The collaborator researchers were advised to estimate the monaural SRT for NH adults and run a comparison between the monaural and binaural SRTs. They used the PAAST to estimate the monaural and binaural SRTs of 10 NH Arabic-speaking adults using the same procedure in the current study. Their data showed normal distribution and the findings are provided in Table 5.3.

Table 5.3 The mean SRT (95% CI) obtained from the right ear, left ear and bilaterally

	Mean SRT (dB SNR)	**SD (dB)**	**95% CI (dB SNR)**
Right ear	−8.5	1.6	−7.9 to −9.1
Left ear	−8.8	1.8	−8.1 to −9.5
Binaural	−10.2	1.1	−9.8 to −10.6

The findings showed that the mean binaural SRT obtained in Study 4 (−10.9 dB SNR) is close to the mean binaural SRT obtained by the subjects in Palestine (−10.2 dB SNR) with a difference of less than 1 dB. Thus, the findings of the Palestine study support the findings of Study 4 and indicate high concurrent validity of the PAAST by showing that the results of other research that replicate the work concur with the results of the current study.

The mean SRT obtained monaurally using the PAAST implemented in the iPad application (right ear= −8.5, left ear= −8.8 dB SNR) is comparable to the findings of previous studies that aimed to validate different SIN tests in different languages. For instance, Smits *et al.* (2013), who validated a Dutch telephone digit-in-noise test, reported a mean monaural SRT of −8.8 dB SNR. Zokoll *et al.* (2012) and Ozimek *et al.* (2009) also reported comparable findings when digit-in-noise tests were developed in German and Polish with a mean SRT of −9.3 and −9.4 dB SNR respectively. Plomp and Mimpen (1979) reported a mean SRT of −6.2 dB SNR for the right ear, −5.6 dB SNR for the left ear and −7.3 dB SNR for the binaural testing. The differences between the SRTs reported by Plomp and Mimpen (1979) and the SRTs obtained in Study 4 can be attributed to the difference in the speech materials (sentences vs monosyllabic words). As discussed in Chapter 2, Section 2.3.3.2, words are expected to be easier to identify than sentences. Thus, it is expected to score better in SIN tests that uses words than SIN tests that use sentences. The difference between the mean SRT of the monaural and the binaural testing obtained by the PAAST (1.5 dB) showed general agreement with the difference between the mean SRT of the monaural and the binaural testing reported by Plomp and Mimpen (1979) and Wilson (2003).

In sum, the findings indicate that the SRTs of NH Arabic-speaking adults obtained by the PAAST binaurally and monaurally are reasonable as they are comparable to the SRTs obtained by other SIN tests, although the comparison was made with SIN tests that are not in the Arabic language. The availability of this information about the performance of NH Arabic-speaking adults in SIN testing, which was not available before, is valuable for clinical and research purposes. It could be used as a reference for other Arabic SIN tests that are developed in the future.

5.5.2 Learning effect

A weakness that was noticed in the methodology of this study is that the order of the tests was not randomised (the 3 rounds of the MATLAB testing was completed followed by 3 rounds of the iPad testing). This might be a possible explanation for the variation in the performance of the subjects that were detected in iPad-R1. It is possible that the subjects were tired as they were performing the test for the fourth time. However, if this was true, it was expected for the performance in the R2 and R3 of the iPad testing to show such a variation, but this was not seen. Besides, no statistical difference was detected between the performance of the subjects when they took the test in both conditions (MATLAB and iPad). It is therefore expected that the variation in the performance in iPad-R1 happened by chance.

No statistical differences were found between the performance of the subjects in the different rounds except for R1 and R3 in the MATLAB setting where R3 showed some improvement (0.8 dB SNR). This could be drawn back to a possible learning effect. It was noticed and reported in previous reliability studies that learning effects are not uncommon in repeated measurements. For instance, Smits *et al.* (2013), who assessed the reliability of the Dutch digit-in-noise test, found a significant difference between the performance in R1 and rounds 3, 4 and 5 in the first session (improvement of about 1.2, 1.5 and 1.6 dB in SRTs) and explained that by a learning effect. However, for the second session, which took place on the same day, and the third, fourth and fifth sessions, which took place on subsequent days with a duration ranging from one day to a few weeks after the first session, Smits *et al.* (2013) did not find statistical differences between the performance of the different rounds. The findings of Smits *et al.* (2013) indicate that a learning effect does not necessarily occur consistently each time the subject performs the test, which might explain the occurrence of a probable learning effect in the third round in the current study but not in the following rounds.

Another study that found statistical differences between the SRTs in different rounds is the one by Lovett *et al.* (2013), which reported an improvement in the performance of children who performed the McCormick Toy Discrimination Test (MTT) in noise twice. Some children performed the two rounds of the tests on the same day, whereas others performed the second round in a subsequent session that took place a few days (1-36 days) after the first session. The mean difference of SRTs between the two sessions was 1.6 dB. Sheikh Rashid *et al.* (2017) also reported a learning effect of 0.7 dB when children performed the Dutch online SIN screening test twice on the same day (SRT-R2 better than SRT-R1). Furthermore, learning effects were found by Ousey *et al.* (1989) with the IHR-McCormick automated Toy Test (improvement of 1.2 dB in SRT) and Plomp and Mimpen (1979) with sentence test.

Apparently, learning effects are common when repeating SIN tests more than once. However, with the PAAST, it is not expected for a learning effect to cause uncertainty about the findings clinically. This is because the expected improvement in SRT that would result from a learning effect, if it occurs, would be very small (less than 1 dB), which is not a clinically material mean shift. Besides, the learning effect was not seen until the third round, and it is uncommon in clinics to perform three repetitions of the same test within the same session.

5.5.3 Test-retest reliability

The reliability of the PAAST was assessed using three statistical measurements: the within-subject variation between the SRTs, the repeatability of the SRTs and the correlation coefficient between the SRTs. The PAAST showed good test-retest reliability for all three measurements.

Some previous studies reported smaller SDω (0.9 dB) (Plomp and Mimpen, 1979; Smits *et al.*, 2004) than the SDω found in Study 4 (1.6 dB), however, the difference is just above half a dB. Other studies reported comparable SDω. For example, Summerfield *et al.* (1994) reported a SDω of 1.3 dB which is close to the SDω found with the iPad settings (1.6 dB). Similarly, MacLeod and Summerfield (1990) and Nilsson *et al.* (1994) reported a SDω of 1.2 and 1.1 dB, respectively, for sentence-in-noise tests.

Other measurements that were used to assess the reliability of the PAAST were the repeatability and the correlation coefficient between the SRTs. Repeatability of 4.4 dB was found with the iPad testing, which is comparable to the repeatability of the MTT when performed by adults in noise (3.6 dB) and better than the repeatability of the MTT when performed by adults in quiet (6.3 dB) (Summerfield *et al.*, 1994). In addition, the PAAST showed a highly positive correlation coefficient between the SRTs across the rounds (ICC: r= 0.7), which is similar to the correlation coefficient of the MTT in noise as reported by Lovett *et al.* (2013) (r= 0.7) and better than that reported by Lovett *et al.* (2012) (r= 0.04). A better correlation coefficient was reported by Summerfield *et al.* (1994) for the MTT in quiet (r= 0.95), which might be caused by the large sample size included in that study (n= 122).

Some studies that assessed the reliability of speech perception tests used the steepness of the slope of the psychometric function of the speech stimuli to determine the reliability of the test. Recall from Chapter 4, Section 4.4.1.1, that steeper slopes lead to a more accurate estimation of the SRT because, with a small increase in the stimulus level, the performance levels increased largely. The mean slope of the psychometric function of the words of the PAAST found in Study 3 was 15%/dB (see Chapter 4, Study 3). This is favourably comparable to the mean slope reported by Yuen *et al.* (2019) (12.5%/dB) and Smits *et al.* (2013) (5.8%/dB when CVC words were used). It is close to the mean slope reported by Plomp and Mimpen (1979) (15%/dB), Smits *et al.* (2004) (16%/dB), Smits *et al.* (2013) (17%/dB when digits were used) and Zokoll *et al.* (2012) (17.6%/dB). This indicates that the reliability of the PAAST is comparable to the reliability of other SIN tests.

In sum, the PAAST showed good test-retest reliability with NH adults when the test was administered for three rounds. It provides reliable results for clinical purposes as its reliability measurements are comparable at least to the MTT, which has the same test principle.

5.5.4 Test duration

Administering the test through the iPad required a longer time than through the MATLAB settings (4.6 and 2.1 minutes respectively). This is not surprising since the iPad interface was developed to be child-friendly, and for that reason, a screen with a smiley face pops up after each response and lasts for 2 seconds (Figure 5.5). The regular appearance of the smiley-face screen is expected to be the reason behind the increase in the iPad's testing time. However, visual enforcements are important to attract the attention of young children and to maintain it for a longer period (Elliott *et al.*, 1979). A duration of around 5 minutes is thought to be acceptable clinically. However, the test duration was also assessed with children in Study 5 (see Chapter 6) in order to find the appropriateness of the duration of the PAAST in testing children.

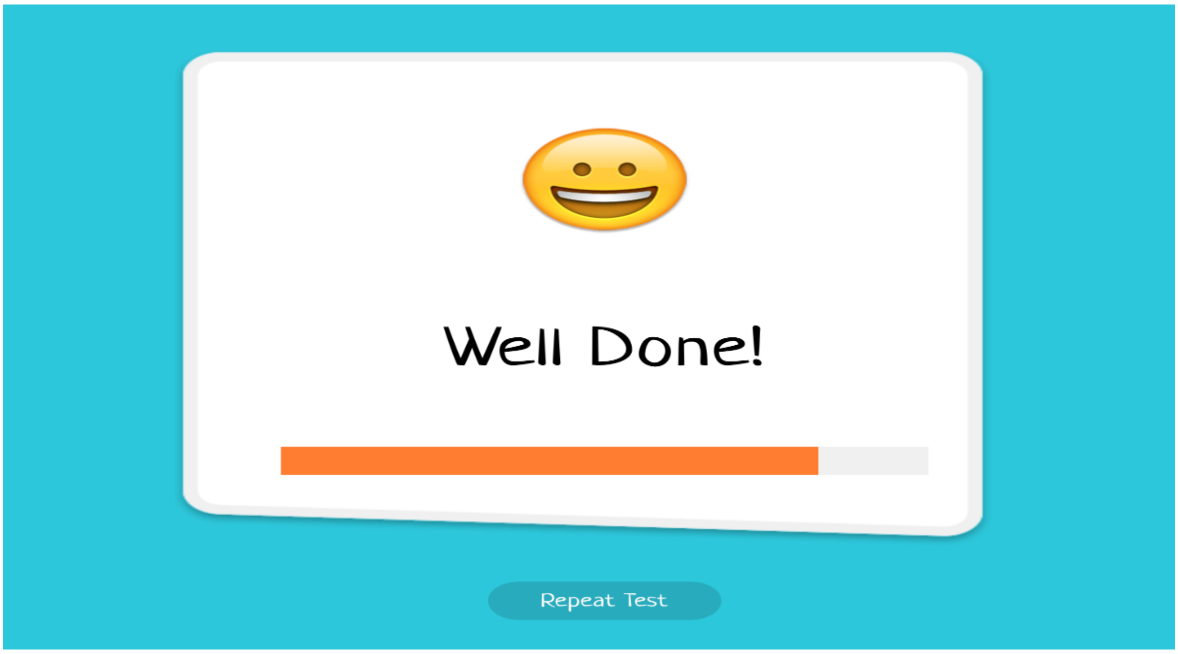


Figure 5.5 Screenshot of the visual reinforcement used in the PAAST iPad application (a smiley face)

5.6 Conclusions

The current study aimed to provide data about the performance of NH Arabic-speaking adults with the PAAST, to assess the test-retest reliability of the PAAST implemented in an iPad application against its reliability when performed through MATLAB settings, and to estimate the test duration. Adults were tested for that purpose to ensure that using the iPad application provided reliable results prior to proceeding with further investigations with children.

The findings of Study 4 are summarised as follows:

- performing the PAAST by NH adults through the iPad leads to similar results as if it is performed through the MATLAB
- the mean SRT in noise with the PAAST for NH adults is −11.2 dB SNR
- when administered through the iPad, the PAAST can predict an SRT that lies within ±3.1 dB of an adult true threshold
- a difference of >4.4 dB between two SRTs obtained by the iPad in two different conditions (e.g. aided and unaided) is considered a true significant difference.

Chapter 6 Studies 5 and 6: Evaluating the PAAST with normal-hearing children and children with SNHL

6.1 Overview of the chapter

In Study 4, the test-retest reliability of the Paediatric Arabic Auditory Speech Test (PAAST) implemented in an iPad application was assessed with normal-hearing (NH) adults and compared against its reliability when performed through a MATLAB software using tried-and-tested code (considering the MATLAB version as an ideal). The findings revealed that the PAAST implemented in the iPad application produces reliable results. However, Study 4 assessed the performance of adults whereas the PAAST mainly targets children. Therefore, it was still important to assess how the PAAST worked with children (the target population); both NH children and children with sensorineural hearing loss (SNHL).

This chapter consists of 2 studies that assessed the performance of Arabic-speaking children in speech-in-noise (SIN) testing. While Study 5 investigated the performance of NH children, Study 6 investigated the performance of children with SNHL when tested by the PAAST. The motivations and aims of Studies 5 and 6 are discussed in the introduction of this chapter.

6.2 Introduction

The primary motivation for the development of the PAAST was to develop a hearing screening tool that is sensitive to SNHL. Because children with SNHL have difficulty understanding SIN, a decision was made to develop an Arabic SIN test for children, particularly as there is a lack of paediatric SIN tests for Arabic-speaking children. However, other gaps of knowledge were noticed, and the aim was to address these in Studies 5 and 6 (some of the gaps are general and some are specific to the PAAST). **General gaps of knowledge** were as follows. **First**, there is no information in the literature about the performance of Arabic-speaking children in SIN testing. Although one would expect the performance of Arabic-speaking children to follow the same pattern of children speaking other languages, there is no research evidence of this. Therefore, it was of interest to investigate the performance of Arabic-speaking children in SIN testing. Providing such information fills the gap in knowledge; besides, it could be used as a baseline for future research work. For instance, it could be used to validate the outcomes of other Arabic SIN tests that are developed in the future.

Second, there is no information in the literature about the effect of age on the performance of Arabic-speaking children in SIN testing. Again, although it was expected to find that the age has an effect on Arabic-speaking children's SRTs in noise as documented for children speaking other languages such as English, Dutch, French and Mandarin (Fallon *et al.*, 2000; Johnson, 2000; Litovsky, 2005; Stuart, 2005; Johnstone and Litovsky, 2006; Stuart *et al.*, 2006; Talarico *et al.*, 2007; Stuart, 2008; Vaillancourt *et al.*, 2008; Zheng *et al.*, 2009; Lovett *et al.*, 2012; Sheikh Rashid *et al.*, 2017; Vickers *et al.*, 2018), there is no evidence-based information about this in the literature. It was therefore considered that this research would explore this topic to provide a scientific evidence base. Also, investigating the effect of age on children's SRT in noise would help in providing normative data and determining cut-off points between the performance of NH and hearing-impaired children that is age-appropriate for children. To clarify, if no age effect was found, the same cut-off value could be used for children of different ages. In contrast, if an age effect was found, more than one cut-off value should be determined (or a correction factor should be used) for each year of age depending on the amount of the effect.

Third, there is no information in the literature about the performance of Arabic-speaking children with SNHL in SIN testing. Providing such information is of equal importance as providing information about the performance of NH children. It would act as guidance for audiologists as to what to expect when testing children with SNHL. **Fourth**, it was unknown whether there is a relationship between PTA thresholds and SRT in noise for Arabic-speaking children. As discussed in Chapter 2, Section 2.3.1, a positive correlation was reported in literature between PTA thresholds and SRT in noise with a variation in the strength of the correlation between different studies (Festen and Plomp, 1983; Glasberg and Moore, 1988; Smoorenburg, 1992; Kramer *et al.*, 1996; Beattie *et al.*, 1997; Divenyi and Haupt, 1997; Hagerman, 2002; Smits *et al.*, 2004; Denys *et al.*, 2018). Again, it is unknown whether the case is the same for Arabic-speaking listeners. Therefore, it was of interest to investigate the correlation between PTA thresholds and SRT in noise.

Other **gaps that are specific to the PAAST** were as follows. **First**, although the words of the PAAST were selected carefully to be familiar to young Arabic-speaking children, it was unknown whether this is true particularly for hearing-impaired children whose language development is expected to be delayed as a result of the HL. Recognising any difficult words at this stage would be useful in order to take appropriate action (e.g. replace the pair), before releasing the PAAST for public use. **Second**, the duration required by children to complete the PAAST was also unknown. It is necessary to estimate the average duration in order to be able to tell whether it is appropriate to use the PAAST to test children and whether it could be used practically for screening purposes and in clinics for diagnostic and rehabilitative purposes. This is because it is not expected for children, particularly of a young age, to engage in a test for a relatively long time due to their expected short attention span in comparison to adults. Besides, it is preferable in clinical situations to have tests that could be completed in short durations in order to account for the clinics' workload. Additionally, if the PAAST to be used for children's hearing screening, it should meet the criteria of screening tests, which includes being a quick test. For example, if it appeared that the average duration required by young children to complete the PAAST is more than 10 minutes, then the procedure of administering the test should be revised to reduce the test's duration as much as possible without affecting the accuracy of the test. To address this issue, the duration of completing the PAAST was assessed in Study 5 with NH children (in controlled settings done by the researcher) and in Study 6 with hearing-impaired children (in clinical settings done by clinicians).

Third, the test-retest reliability of the PAAST when performed by NH children was unknown. Although the PAAST showed high reliability when performed by adults, children's performance might show a variation (both between and within-subjects). If the PAAST showed poor test-retest reliability with NH children, a revision to the test procedure, the selected speech stimuli etc. should be made before proceeding with further experiments with hearing-impaired children. Therefore, the reliability of the PAAST with NH children was assessed in Study 5.

Fourth, it was unknown whether the PAAST is sensitive to SNHL or not. Recall that the target population in this PhD project was children with SNHL. The PAAST was developed with an ultimate aim to improve the services provided to those children in terms of early identification and other diagnostic/post-diagnostic purposes. In order to ensure that the development of the PAAST would help in achieving this aim, it was essential to investigate the ability of the PAAST to identify children with SNHL. The aim was to assess the ability of the PAAST to identify children with SNHL based on the distortion component of HL rather than the audibility component. This is because children with SNHL suffer more from Class-D HL, which interferes primarily with their ability to understand SIN (the classifications of HL as given by Plomp (1978) were discussed previously in Chapter 2, Section 2.2.1). Therefore, the ability of the PAAST to identify those children was assessed by emphasising the distortion component and reducing the audibility component as much as possible (see Section 6.6.2 for details).

The ability of the test to identify children with SNHL in terms of Class-D HL and to differentiate between them and NH children was investigated in Study 6. It is worth reminding the reader here that HL of Class D is characterised by distortion to the received stimuli as a result of impaired psychoacoustic aspects of hearing such as the frequency selectivity, temporal resolution, etc. This class of HL as described by Plomp (1978) decreases the ability to understand SIN because it interferes with the SNR. To address the gaps of knowledge mentioned above, the following aims were set for Studies 5 and 6.

6.2.1 Aims

- **Study 5 (NH children) aimed to:**
    1. assess the familiarity of the words of the PAAST in NH children
    2. estimate the duration required by NH children to complete the PAAST
    3. provide data about average SRT in noise for Arabic-speaking children
    4. assess the test-retest reliability of the PAAST with NH children
    5. investigate the effect of age on SRT in noise of NH Arabic-speaking children
    6. provide preliminary normative data for children.

- **Study 6 (children with SNHL) aimed to:**
    1. assess the familiarity of the words of the PAAST in children with SNHL
    2. estimate the duration required by children with SNHL to complete the PAAST in clinical settings
    3. investigate the ability of the PAAST to identify children with SNHL based on Class-D HL and to differentiate between them and NH children
    4. investigate the relationship between PTA thresholds and SRT in noise for Arabic-speaking children with SNHL (particularly for the PAAST when the distortion component of SNHL was emphasised).

6.3 Study 5 (NH children): methods

Arabic-speaking children were recruited in Riyadh, KSA for the purposes of the study. The inclusion criteria were: (1) children aged 3.0-12.0 years, (2) native Arabic speakers, (3) PTA thresholds of 20 dB HL or less for the frequency range from 0.25-8 kHz, (4) normal language development, (5) no vision problems, and (5) free of learning disabilities. Parents were asked to fill in an otological-health questionnaire about the child to rule out HL-related symptoms including recent ear disease, ear surgery, tinnitus, recent exposure to loud noise and family history of HL and to rule out other disabilities such as vision and learning disabilities. Children were allocated into 2 groups according to their ages, (Group 1 included children aged 3.0 years to 5 years 11 months and Group 2 included children aged 6.0-12.0 years). Children were distributed into 2 groups in order to look for an age effect on the SRTs of the children and to provide normative data for each group, if necessary. Written consent was obtained from parents of all children. Additionally, written assent was obtained from the older children, and verbal agreement was obtained from the younger children prior to performing any testing.

6.3.1 Sample size

The sample size necessary to assess the test-retest reliability was calculated as follows. It was expected to observe more variation in the performance of children than adults. In Study 4, the maximum mean difference was between round 1 (R1) and round 3 (R3) (0.8 dB SNR) and the maximum standard deviation (SD) was in R1 (2 dB SNR). It was expected for the mean difference between the rounds to be lager when children performed the test. Therefore, the sample size was calculated based on an expectation of a mean difference of 1.3 dB SNR (half a dB more than adults) and a SD of 2 dB SNR. Based on this, a minimum of 19 subjects was required to be tested in each group in order to detect a mean difference of 1.3 dB SNR with a power of 0.8 and alpha 0.05 (Howell, 1997).

The current study aimed to provide *preliminary* normative data, not large-scale normative data because at this stage, the sensitivity of the PAAST to SNHL was not known. It was believed that it was not reasonable to recruit a large number of children and then find at a later stage that the test was not even sensitive to SNHL. Therefore, the same sample was used to provide preliminary normative data. Similarly, the effect of age was investigated *primarily* on the same number of children. However, if an age effect was detected in this study and the PAAST was found sensitive to SNHL, a large-scale normative data study could be done in a future study including an appropriate number of children at each age group to provide normative data and correction factors for each year of age.

6.3.2 Materials and calibration

- Portable Amplivox audiometer (serial number: 31557), used to screen the subjects' hearing
- Sennheiser HD 200 headphones
- Norsonic sound level meter (SLM) Nor123 (serial number: 1323063), used to monitor the background noise
- An iPad with the developed PAAST application
- YAMAHA mixing console with two microphone inputs, model MG06 (serial number: BGXJ01035). This mixer was used to process the output from the iPad to the headphones.

Objective calibration was conducted before carrying out any data collection. The level of the noise was calibrated using a SLM, manufactured by Brüel & Kjær (type: 2260, serial number: 1853816), a sound level calibrator (94 dB, 1 kHz) manufactured by Brüel & Kjær (type: 4230, serial number: 861555) and an artificial ear (type: 4153, serial number: 198145).

6.3.3 Procedure

Each child attended one session lasting around one hour. The child was seated on a comfortable chair in a quiet room. The background noise in the room was monitored using the SLM. The maximum background noise did not exceed 40 dB A, which is within the maximum acceptable ambient noise level in the testing room when screening for air-conduction thresholds (Gelfand, 2009). At the beginning of the session, the child was screened for HL monaurally using pure tone audiometry (PTA) for the frequency range from 0.25-8 kHz at 20 dB HL. The child was informed that s/he would hear a tone and was asked to raise his/her hand each time s/he heard the tone even if it was very low. Children who passed the screening at 20 dB HL were included.

Afterwards, the child was tested with the PAAST implemented in an iPad application using the same adaptive procedure described in Chapter 5, Section 5.3.1. The iPad was placed on a small table in front of the child. The tester started by naming the pictures that represented the test words and asked the child to point to the picture that was named. This was to look for the familiarity of the words and to ensure that the words were within the vocabulary of the child. In cases where the child did not recognise any of the words, that word and its pair were excluded from the test. Afterwards, the child was informed that the headphones would be placed over his/her ears, through which s/he would be listening to sentences in the form of "Where is the (word)?" and that s/he was asked to touch the picture of the word that s/he heard. The child was asked to guess if s/he was not sure what the word was. Children were tested monaurally (the right ear was tested for half of the children and the left ear was tested for the other half randomly in each group). The aim was for each child in each group to perform the test 3 times in order to explore the test-retest reliability of the PAAST. The children were offered 5-10 minutes break between the tests.

Ethics approval was provided by the ethics committee at the University of Southampton in the UK (identification number 40367).

6.4 Study 5 (NH children): results

Descriptive data of the subjects are shown in Table 6.1.

Table 6.1 Descriptive data of NH children

	Mean age (years)	**Gender**		
	SD (years)	**Male**	**Female**	**Total**
Group 1 (young children)	4.7 (SD= 0.9)	12	8	20
Group 2 (older children)	9.2 (SD=1.8)	10	10	20
Complete sample	6.9 (SD=2.7)	22	18	40

6.4.1 The familiarity of the children with the words of the PAAST

All the words of the PAAST were found to be familiar and within the vocabulary of 93% of the children (37 children). Three children aged 3.0, 3.0 and 3.5 years failed to recognise some of the words (5 words), which led to excluding a pair of words for 1 of these children and 2 pairs for the other 2 children. However, the unrecognised words were different for the 3 children (i.e. no common words were found to be unfamiliar to them). The words were: “elephant, worms, rice, banana and roaster”. Interestingly, one non-Saudi child (from Sudan) aged 5 years was familiar with all the PAAST words and was able to perform the test.

6.4.2 Test duration

The mean test duration was 5.1 minutes (SD=0.5 minutes) for Group 1 and 4.6 minutes (SD= 0.8 minutes) for Group 2. The two sets of data did not show a statistically normal distribution (Shapiro-Wilk test: p= 0.01 and p= 0.04 for the younger and older children, respectively).; therefore, the parametric statistical test was conducted. Although a statistical difference between the duration required to complete the test by the two groups was found (Mann-Whitney test: U= 100.5, p=0.007), the mean difference was very small (0.5 minute).

6.4.3 Mean SRT of NH children

The aim was for each child to perform the test for 3 rounds. However, all 20 children in Group 1 (young children) failed to perform R3 due to tiredness and/or unwillingness with 3 children only completing R1. In Group 2 (older children), 6 children failed to perform R3 for the same reasons.

The mean SRT across the 3 rounds was −4.9 dB SNR (SD= 2.4 dB, 95% CI: −5.7 to −4.1 dB) for Group 1 and −7.7 dB SNR (SD= 2.2 dB, 95% CI: −8.4 to −6.9 dB) for Group 2. The mean SRTs of each round obtained by the children of both groups are provided in Table 6.2. The boxplots in Figure 6.1 also show the SRTs obtained by the children of both groups across each round.

Table 6.2 Mean SRT across each round obtained by both groups of NH children

		Number of children	**Mean SRT (dB SNR)**	**SD (dB)**	**95% CI of mean (dB)**
Group-1 (young children)	**R1**	20	−4.9	2.1	−5.9 to −3.9
	R2	17	−4.7	2.7	−6.2 to −3.4
	Average of R1 & R2	17	−4.9	2.4	−5.7 to −4.1
Group-2 (older children)	**R1**	20	−7.7	2.2	−8.8 to −6.6
	R2	20	−7.4	2.1	−8.3 to −6.4
	R3	14	−8.2	2.6	−5.7 to −4.1
	Average of all 3 rounds	14	−7.7	2.2	−8.4 to −6.9

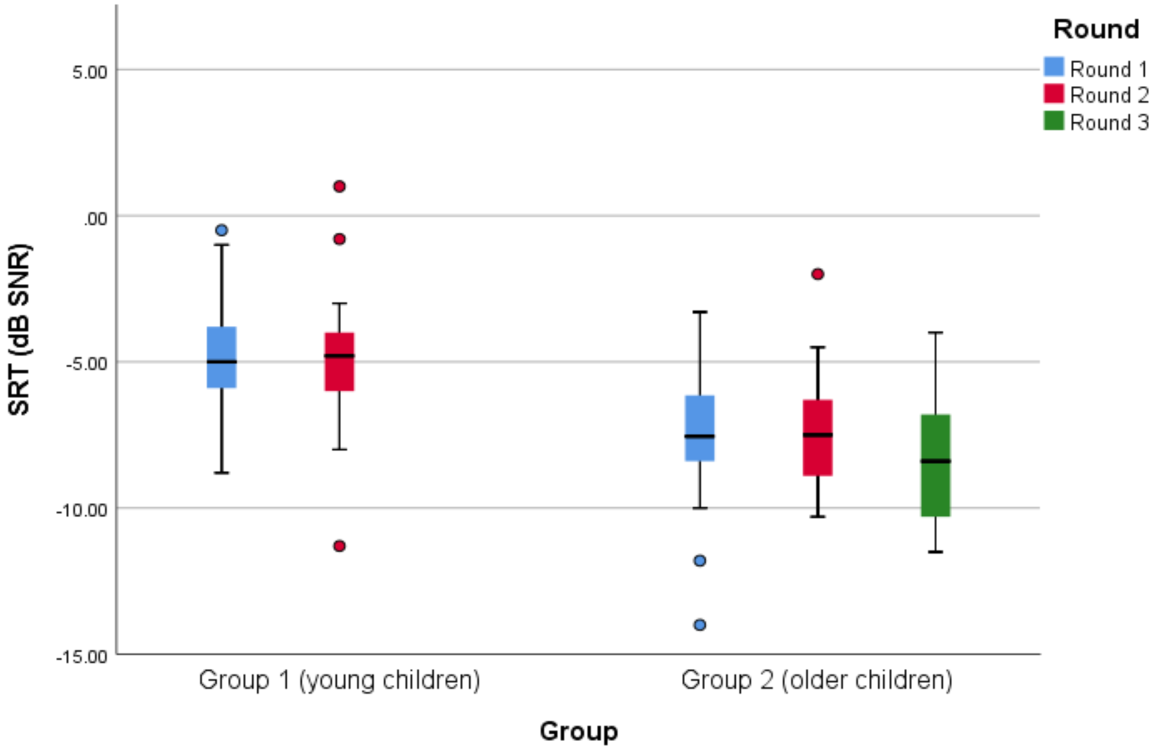


Figure 6.1 Boxplots representing the SRTs across each round for the two groups of NH children

It can be seen in Figure 6.1 that the variation in the performance of the children of both groups in each round is small. Additionally, the performance of the children of each group across the rounds looks similar with only small variations. However, it was noticed that the performance of the older children in R3 looked relatively better than R1 and R2, indicating a possible learning effect. Four outliers were detected in Group 1, 1 of them in R1 and 3 in R2. Another 3 outliers were detected in Group 2, 2 of them in R1 and 1 in R2. However, it is reassuring that the outliers present both above and below the box plots, indicating that the outliers were not consistently performing better or worse than the other children. When performing the statistical analysis, the outliers were not excluded in order to not bias the results.

6.4.4 The test-retest reliability and the effect of age on the SRTs of the children

In order to look for a systematic variation in the performance of the children of both groups, the following statistical analysis was carried out. For Group 1, the paired sample t-test was carried out to look for a difference in the mean performance of the children between the 2 rounds since the data for the 2 rounds were at least approximately normally distributed (Shapiro-Wilk test: p= 0.5 & 0.4 for R1 and R2 respectively). The test revealed no statistically significant difference between the means of the 2 rounds (t_{16}= −0.8, p= 0.4).

To look for a difference in the mean performance of the children in Group 2 across all rounds, the one-way repeated measures ANOVA was carried out. The data of the 3 rounds for Group-2 were at least approximately normally distributed (Shapiro-Wilk test: p=0.4, 0.5 & 0.2 for R1, R2 and R3 respectively) and Mauchly's test of sphericity showed that the assumption of sphericity was not violated [$X^2(2) = 1.2$, p=0.5]. Accordingly, the repeated measures ANOVA with sphericity assumed was calculated and showed no statistically significant main effect of the rounds [F (2, 26) = 1.6, p=0.2].

In addition, the mixed ANOVA was conducted to investigate the impact of age and rounds on the SRTs of children in both groups. In order to do that, R3 that was performed by Group 2 was excluded. Additionally, subjects who did not complete R2 in Group 1 were excluded (3 children). Data for all rounds were approximately normally distributed (Shapiro-Wilk test: p= 0.4, 0.6, 0.5 & 0.4 for Group 2/R1, Group 2/R2, Group 1/R1 and Group 1/R2 respectively). Data for R1 and R2 were shown to be homogenous (Leven's test: p=0.6 and 0.9 for R1 and R2, respectively, and the Box's test of equality of covariance matrices: p=0.5). The mixed ANOVA revealed no significant main effect of the rounds, F (1, 35) = 0.8, p= 0.3 with a mean difference of −0.3 dB (95% CI: −0.4 to −0.5 dB). According to this finding, if there is a learning effect (as expected in R3 for the older children), it would be less than 0.5 dB. The main effect of age was statistically significant, F (1, 35) = 16, p<0.0001), with a mean difference of −2.5 dB (95% CI: −2.4 to −2.6 dB). The older children (Group2: mean SRT= −7.7 dB) showed lower (better) SRT than the younger children (Group 1: mean SRT= −4.9 dB). Additionally, no significant interaction between the age and rounds was found by the mixed ANOVA, F (1, 35) = 0.03, p= 0.8.

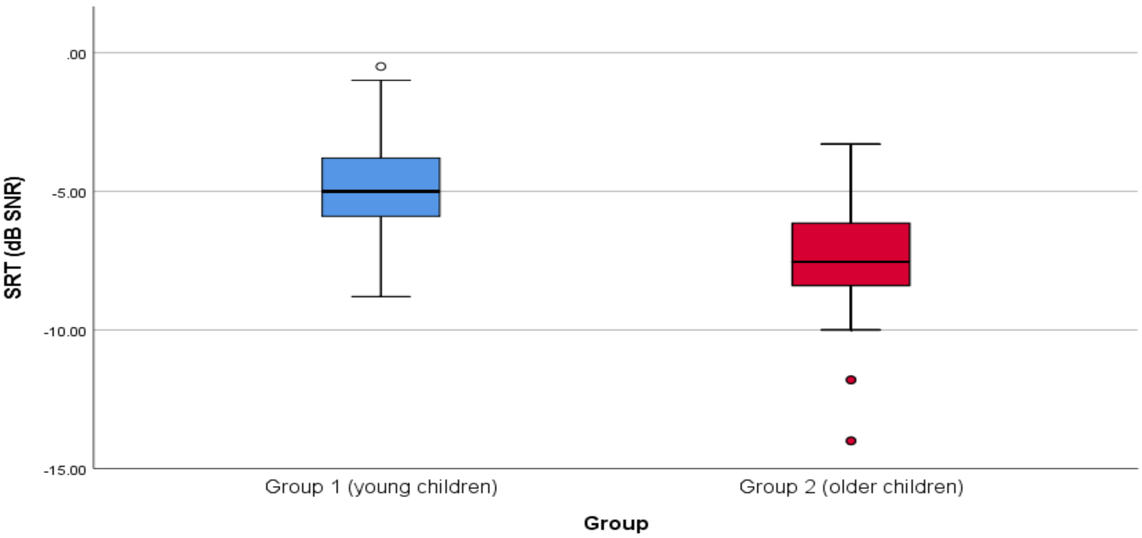


Figure 6.2 Improvement in the SRT with age (R1)

The improvement in the SRTs with age is shown in Figure 6.2. As can be seen, the boxplots showed that the SRTs obtained by the young children were higher (worse) than those obtained by the older children. However, the two groups of children were not controlled for variables other than age; therefore, it was not expected that age would explain 100% of the variation in the performance between the two groups. Besides, the effect of age on the SRTs was expected to be gradual, which is not shown in the boxplots. Therefore, a scatter plot, which represents the SRT of each child in both groups, was created. As it appears in Figure 6.3, there is a negative linear correlation between the age and the SRT: as the age increases, the SRT decreases (improves) gradually.

A linear regression analysis was carried out with the SRT as the dependent variable and the age as the independent variable and suggested a strong correlation between the age and the SRTs (r= 0.7). It also indicates that more than half (55%) of the variation in the performance of the children of the 2 groups is explained by age (R^2= 0.55). The regression equation (SRT= 1.63 - 0.66 age) suggests that between the age of 3 and 12 years, there is a significant improvement of 0.66 in the mean SRT score per year of age (95% CI: −0.8 to −0.4 dB).

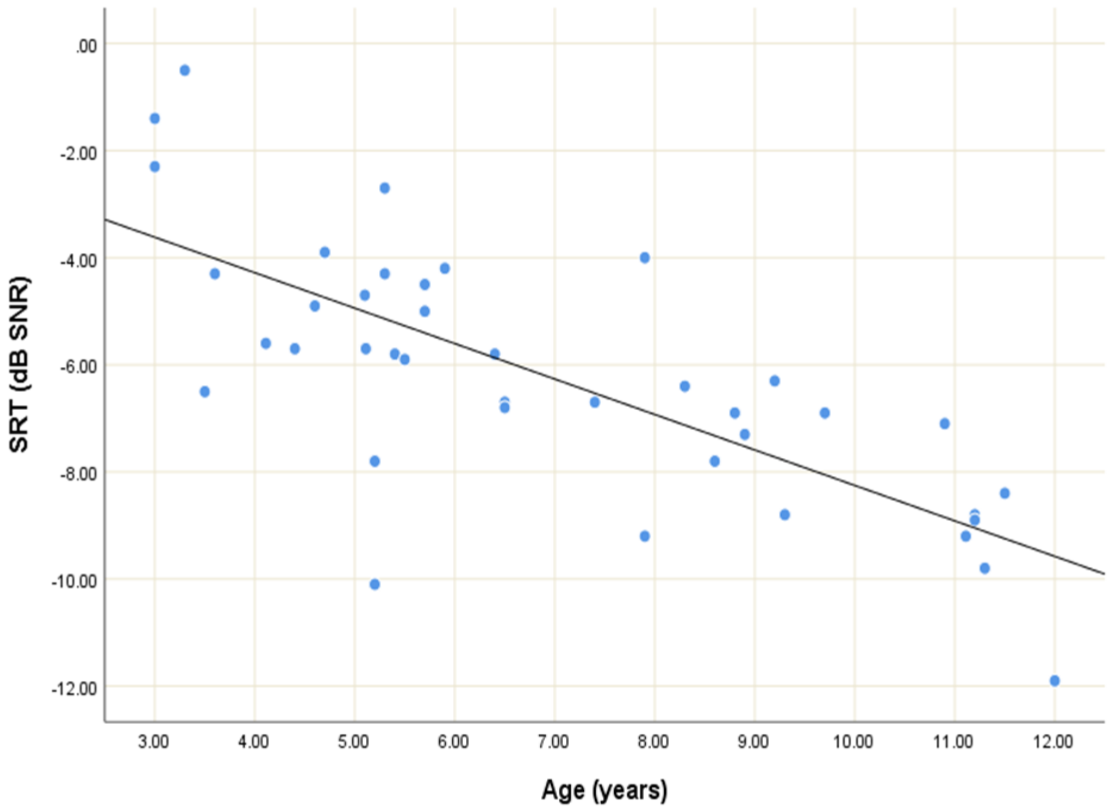


Figure 6.3 Scatter plot showing the correlation between the age and the SRT (R1) of all children (n=40). Each dot represents one child. (R^2 =0.55)

The ANOVA analysis, which was carried out at the beginning of this section suggested that there is no systematic variation in the performance of the children of both groups. The following statistical analysis aimed to look for the random variation in the performance of the children of both groups. First, the **variability** in the SRTs obtained by each child was assessed by calculating the within-subject standard deviation (SDω) using Equation-1 (see Chapter 5, Section 5.3.3.4). The SDω was equal to 1.8 dB and 1.6 dB for Group 1 and Group 2 respectively. It was found that the true SRT score for any one-measurement value lies within ±3.5 dB for Group 1 and ±3.1 dB for Group 2 with a probability ≥0.95. This indicated that the PAAST could predict an SRT that lies within ±3.5 dB of young children's true threshold and ±3.1 dB of older children's true threshold.

Second, the **repeatability** of the test was calculated by multiplying the **SDω by √2 x 1.96** and showed that a difference greater than 4.9 dB and 4.4 dB between any two SRTs obtained in two different conditions (e.g. aided and unaided) for Group 1 and Group 2 respectively is considered a true statistically significant difference.

The **intra-class correlation coefficient (ICC)** statistical test was carried out for each group of children in order to assess the correlation coefficient of the SRTs between the rounds. It was found that the PAAST provides reliable results when repeated a few times with a reliability of r= 0.7 (95% CI: 0.1 to 0.9, p= 0.02) for Group 1 and r= 0.8 (95% CI: 0.4 to 0.9, p= 0.001) for Group 2. The ICC of 0.7 and 0.8 indicates a high positive correlation (Mukaka, 2012), which confirms the reliability of the test in examining children.

6.4.5 The normal range of SRTs

The normal range, sometimes referred to as a reference range, is a range of values that is used as a reference to interpret the results of a given test. Having a score that falls within this range indicates that the subject's score is within normal limits and thus the subject is not expected to have the condition. Determining a normal range is important for clinical practice to differentiate between normal and abnormal results. However, although ideally tests are expected to differentiate 100% between the two populations, in the real world there is always some overlap between them (Whyte and Kelly, 2018).

In the current study, since it was found that age has an effect on the SRTs of the children, it was necessary to find an age-appropriate normal range. In other words, it would not be accurate clinically to come up with one normal range for children of all ages (3.0-12.0 years) because a variation was detected between the performance of the young and the older children. Therefore, two normal ranges were calculated: one for young children (3 years to 5 years 11 months) and one for older children (6.0-12.0 years).

At this stage, it was necessary to decide which mean SRT to use for this purpose (mean of R1, R2 or R3). It was not appropriate to use the mean of the 3 rounds since patients are expected to perform the test only once in clinics. Therefore, it is not appropriate to compare an SRT obtained from a single test to a normal range that is calculated based on an average of multiple tests. Thus, a decision had to be made regarding which mean to use. Since no statistically significant difference was found between the mean SRT of the rounds by the ANOVA test (Section 6.4.4), it was appropriate to use any of them. However, a decision was made to use the mean SRT of R1 because most patients are expected to perform the test once either for screening or diagnostic purposes and most likely that will be their first time performing the test. Based on this, the mean SRT of R1 was used to calculate the normal range for each group.

The normal range was calculated as the **mean ±2SD** (Bland, 2015). The 95% CI of the upper and lower limits of the normal range were also calculated in order to determine how accurate it is to estimate the population's normal range from the current sample. The 95% CIs of the upper and lower limits of the normal range were calculated as follows. First, the standard error (SE) of the limit of the normal range was calculated using the following equation (Bland, 2015):

$$\sqrt{SD^2(\frac{1}{n}) + (\frac{2}{n-1})}$$

where *SD* is the standard deviation of the mean SRT and *n* is the number of subjects.

After that, the 95% CI of the normal range was calculated as the (upper limit of the normal range ±1.96 x SE) to the (lower limit of the normal range ±1.96 x SE) (Bland, 2015). The normal range and 95% CI of the upper and lower limits of the normal range for both groups of children are shown in Table 6.3.

Table 6.3 The normal range of SRT and 95% CI of the upper and lower limits of the normal range for the two groups of children (young and older children)

	Normal range (dB SNR)	95% CI of the lower limit (dB)	95% CI of the upper limit (dB)
Group 1 (young children)	−9.1 to −0.7	−10.3 to −7.9	−1.9 to +0.5
Group 2 (older children)	−12.1 to −3.3	−13.3 to −10.9	−4.5 to −2.2

Based on the calculated normal ranges, the cut-off values were proposed to be −0.7 and −3.3 dB SNR for younger and older children, respectively (which is the higher value of the normal range). Because of the observed effect of age on the SRTs of the children, it might be more appropriate to determine an age-specific cut-off value or a correction factor per year of age rather than a cut-off value for an age range. However, in the current study, it was not possible to determine age-specific cut-off values per year of age because of the small sample size, which resulted in a limited number of children in each year of age, as can be seen in Table 6.4.

Table 6.4 Number of children per year of age

Age	Number of children
3.0 to 3 years 11 months	5
4.0 to 4 years 11 months	4
5.0 to 5 years 11 months	11
6.0 to 6 years 11 months	3
7.0 to 7 years 11 months	3
8.0 to 8 years 11 months	4
9.0 to 9 years 11 months	3
10.0 to 10 years 11 months	1
11.0 to 12.0 years	6

6.5 Study 5 (NH children): discussion

The PAAST was developed originally to target children; thus it was important to investigate the reliability of the test when performed by NH children. The reliability was assessed by comparing the performance of the participating children in more than one round of the test and looking for significant differences between them. Additionally, preliminary normative data were provided for NH children. These data could serve as a reference to differentiate between NH children and children with HL. The subsequent sections discuss the findings of Study 5 in terms of the performance of the NH children, the effect of age on the children's performance, the reliability of measurements of the test and the proposed normal range for NH children. The familiarity of children with the words of the PAAST and the test's duration are discussed at the end of the chapter with the discussion of the findings of the hearing-impaired children (see Section 6.8.1).

6.5.1 Effect of age on the children's performance

It is well documented in literature that the performance of children in speech perception in noise improves with age (Fallon *et al.*, 2000; Johnson, 2000; Litovsky, 2005; Stuart, 2005; Johnstone and Litovsky, 2006; Stuart *et al.*, 2006; Talarico *et al.*, 2007; Stuart, 2008; Vaillancourt *et al.*, 2008; Lovett *et al.*, 2012; Sheikh Rashid *et al.*, 2017; Vickers *et al.*, 2018) and does not reach adult-like levels until adolescence, depending on the SIN listening condition. For instance, the performance of children in speech-shaped noise reaches adult-like levels earlier than their performance in other types of noise such as the babble noise (Corbin *et al.*, 2016).

The difference between the performance of children and adults in SIN tasks is explained mainly by the continuing maturation of the auditory perceptual abilities over the first 10-12 years of age (Elliott, 1979; Elliott *et al.*, 1979; Eisenberg *et al.*, 2000; Fallon *et al.*, 2000; Hartley *et al.*, 2000; Stuart, 2005; Stuart *et al.*, 2006). The continuing development of temporal resolution abilities allows older children to discriminate better between phonemes and thus to recognise words better than young children (Hartley *et al.*, 2000; Stuart, 2005; Stuart *et al.*, 2006). The continuing maturation of the auditory cortex until adolescence (Paus *et al.*, 1999) causes improvement in several functions such as phoneme recognition, speech perception in noise and the use of linguistic cues (Boothroyd, 1997). In general, children require higher SNRs than adults to understand SIN because they are unable to use all the sensory information (Eisenberg *et al.*, 2000).

Other non-sensory factors such as linguistic skills and cognitive abilities were also reported to associate with the performance of children in understanding SIN. The experience with language during childhood results in reorganisation in brain structures (Rauschecker, 1999), which emphasise the improvement of speech perception in noise with increasing language experience. The development in language skills allows older children to utilise their knowledge of language to fill in the gaps and correctly guess the target word from a partially received message (Gat and Keith, 1978; Elliott *et al.*, 1979). The familiarity of the older children with the target words and the increase in the frequency of word usage with age would cause this improvement in their ability to identify the target word (Elliott *et al.*, 1979; Gelfand, 2009). For instance, the word "dog" is familiar to 3-year-old children, but it is expected that 10-year-old children have been exposed more to the word "dog", thus the frequency of words usage might influence their ability to correctly identify the word in challenging listening situations. Cognitive factors such as working memory and attention were also noted to play a role in understanding SIN. For example, children might find it more difficult to recall a target word from a presented sentence than adults (Elliott, 1979). Similarly, inattention is a possible reason for the differences between the performance of children and adults in SIN tasks (Wightman and Allen, 1992). The contributions of cognitive and linguistic factors would raise the SNRs required for children to simulate adult-like performance. However, it is important to note that non-sensory factors influence the ability to understand speech in different listening conditions including quiet conditions. It is therefore expected that these factors might contribute to the worst performance of children in SIN tasks but are not expected to affect their performance mainly.

Study 5 looked at the effect of age on the performance of Arabic-speaking children in SIN tasks. Although it was expected that the performance of Arabic-speaking children in SIN tasks would follow the same pattern of the performance of children speaking other languages and show a developmental age effect, there is no research evidence of this. Exploring this topic to provide a scientific evidence base was therefore considered in this research.

The participating children were allocated into two groups according to their age in order to look for the effect of age on the children's performance. The findings suggested an age effect on the SRTs of the children as the younger children obtained poorer SRTs in comparison to the older children. The finding is consistent with the literature, where children speaking other languages such as English, French, Dutch and Mandarin showed improvement in their performance in SIN tests with age (Fallon *et al.*, 2000; Johnson, 2000; Litovsky, 2005; Stuart, 2005; Johnstone and Litovsky, 2006; Stuart *et al.*, 2006; Talarico *et al.*, 2007; Stuart, 2008; Vaillancourt *et al.*, 2008; Lovett *et al.*, 2012; Sheikh Rashid *et al.*, 2017; Vickers *et al.*, 2018). This indicates that the performance of Arabic-speaking children in SIN testing improves with age just as the performance of children speaking other languages does.

It is acknowledged that the participating children in each group were only matched for age but not for other factors, such as socioeconomic status. The statistical analysis revealed that 55% of the variation in the performance of the children of the two groups is explained by age. Thus, it could not be said that age fully explains the observed difference. However, this was the case in many other previous research studies where cross-sectional rather than longitudinal studies were used which did not mention matching the groups of children for factors other than age (Litovsky, 2005; Stuart, 2005; Johnstone and Litovsky, 2006; Stuart *et al.*, 2006; Stuart, 2008; Vaillancourt *et al.*, 2008; Lovett *et al.*, 2012; Sheikh Rashid *et al.*, 2017; Vickers *et al.*, 2018).

In this study, young children (3-6 years old) required around 3 dB higher SNRs to achieve comparable performance to the older children (6+ to 12 years old) (70.7% correct responses). Similarly, the older children required around 3 dB higher SNRs to achieve comparable performance to adults (see Chapter 5, Section 5.4.1 for SRT of NH adults). Thus, young children required around 6 dB more to achieve comparable performance to adults. This is comparable to the findings of Fallon *et al.* (2002) who reported that 5-year-old children require 5 dB more SNRs than adults to obtain a comparable performance. However, it is acknowledged here that the comparison might not be accurate, particularly as types of noise and speech materials differ between the two studies. While Study 5 measured the performance on monosyllabic words presented in speech-shaped noise, Fallon *et al.* (2002) measured the performance on low-context sentences presented in babble noise. Sheikh Rashid *et al.* (2017), who measured the performance of children on monosyllabic words presented in speech-shaped noise, reported that 5-year-old children required SNRs that were 1.5 dB more than that of 11-12-year-old children to obtain a comparable performance. This indicates that the difference in the performance of the young and older children in Sheikh Rashid *et al.* (2017) is less than the difference in the performance of the young and older children in Study 5. This could be explained by the fact that the youngest children in the study by Sheikh Rashid *et al.* (2017) were aged 5 years, whereas in Study 5 the youngest children were aged 3 years. Thus, the inclusion of younger children in Study 5 might have increased the difference. Talarico *et al.* (2007) also reported a difference in the performance of children aged 6-16 years in understanding words and non-words presented in speech-shaped noise. Older children outperformed younger children in all conditions by around 2-3 dB SNR, which is comparable to the findings of Study 5.

In sum, it was found in this research that there is a developmental age effect on the performance of Arabic-speaking children in understanding SIN. Direct comparison with previous research was not reasonable due to the differences in the ages of the children participating in different studies, types of speech materials and noise. However, at least generally, the findings seem to be consistent with the findings of previous studies.

6.5.2 Mean SRT of NH children

Ideally, the performance of children with the PAAST should be compared to the performance of Arabic-speaking children performing another Arabic SIN test. However, the only available published SIN in the Arabic language is the Arabic version of the word in noise test (WIN) (Rahman, 2018). It was not appropriate to compare the findings of the current study to Rahman's (2018) findings because the two tests (the PAAST and the WIN tests) differ in their measurements. While the PAAST measures the SRT in noise using an adaptive procedure, which allows for obtaining thresholds as low as −20 dB SNR, the WIN test measures the percentage of correct responses at a fixed SNR, with

0 dB SNR being the lowest SNR. Hence, the comparison between the two tests could not be performed. Therefore, the findings of Study 5 were compared to the findings of published SIN tests that measure children's SRT in noise in other languages using adaptive procedures, words as the speech material and speech-shaped noise as the masker.

Because the development of the PAAST was inspired by the McCormick Toy Discrimination test (MTT), it was of interest to find whether the children's performance with PAAST was similar to the performance of children who performed the MTT. Lovett *et al.* (2012) used the MTT to measure the SRT in noise of NH children aged 3.0 to 7.9 years in free-field (the speech was presented from 0°, and the noise was presented from −90°, 0°, +90°). Finding some variations in the children's performance in the two tests was expected because there was a difference in the presentation mode of the speech materials and noise between the two studies, while headphones were used in Study 5, free-field testing was performed in Lovett *et al.* (2012), and the languages of the two tests (Arabic vs English). However, the comparison was still made to roughly judge whether the SRTs obtained from the children who performed the PAAST were reasonable. The mean SRT obtained by the young children in Study 5 (−4.9 dB SNR) is comparable to the mean SRT obtained by the young children in Lovett *et al.* (2012) (around −4 dB SNR was found for children aged 2-5 years when both the noise and the speech stimuli were presented from 0°). The mean SRT obtained from older children (aged 6.0-7.0 years) in Lovett *et al.* (2012) was around −5 dB SNR, which is higher (worse) than the SRT obtained by the older children in Study 5 (−7.7 dB SNR). This could be explained by the inclusion of older children (8.0-12.0 years) in Study 5, thus the better SRTs obtained from those older children might contribute to a better mean SRT for this group. By looking at the SRTs of children aged 6.0-7.0 years (n= 6 children) in Study 5, it was found that the mean SRT was −6 dB SNR, which is comparable to the mean SRT obtained from children of the same age in Lovett *et al.* (2012). Hence, it could be said that the mean SRT obtained from children in Study 5 is comparable to the performance of children of similar ages who performed the MTT in noise, bearing in mind the differences between the two tests.

The findings of Denys *et al.* (2018), who assessed the SRT in noise for children aged 9.0-12.0 years using the digit triplet test in Flemish, also showed comparable findings. Because the children in the study by Denys *et al.* (2018) were aged 9.0-12.0, their SRTs were compared to the SRTs obtained from the older children in Study 5. Higher SRTs (around −9.8 dB SNR) were reported in Denys *et al.* (2018) than the SRTs obtained by the older children in Study 5 (−7.7 dB SNR). However, by looking at the children aged 9.0-12.0 years (n= 10) in Study 5, it was found that the mean SRT is −8.4 dB

SNR, which is comparable to the mean SRT obtained from children of the same age in Denys *et al.* (2018).

Other studies reported different findings for NH children, some of which reported higher (worse) SRTs and some reported lower (better) SRTs than found in Study 5. For instance, Corbin *et al.* (2016), who developed an open-set word recognition test for children in English, reported higher mean SRTs when speech-shaped noise was used. For children aged 5-7 years, the reported mean SRT was −0.5 dB SNR (SD= 2.3 dB), and for children aged 8-12 years, the reported mean SRT was −1.5 dB SNR (SD= 2.2 dB). Apparently, the mean SRTs found in the study by Corbin *et al.* (2016) study were higher than the mean SRT found in Study 5. It is also higher than the mean SRT reported by Lovett *et al.* (2012) although both tests were in English, used the same types of speech stimuli and masker (i.e. words and speech-shaped noise), and performed free-field testing. A possible explanation for the observed differences between the mean SRTs obtained by the children in Corbin *et al.* (2016) and Lovett *et al.* (2012) is that Corbin *et al.* (2016) used open-set speech stimuli consisting of 837 words, which is expected to be harder to predict than the closed-set words used in Lovett *et al.* (2012). The same explanation applied to the differences between the findings of Corbin *et al.* (2016) and Study 5. In addition, the differences in the languages of the tests and the presentation modes (free-field vs headphones) might contribute to the observed differences.

In contrast, Sheikh Rashid *et al.* (2017), who evaluated an internet-based SIN screening test for children in Dutch, reported age-specific SRTs that were much lower than the SRTs found in Study 5. The SRTs in the study by Sheikh Rashid *et al.* (2017) ranged from −13.1 to −15.2 dB SNR for each year of age from 5-10 years old. It is worth noting that Sheikh Rashid *et al.* (2017) measured the SRT binaurally, which was expected to show better performance than the monoaural testing. However, the binaural testing itself was not expected to result in such a large difference between the mean SRT reported by them and the mean SRT found in Study 5 (a difference of around 8 dB for young children).

Similarly, Yuen *et al.* (2019), who assessed young children aged 3.5-5.0 years using the Mandarin spoken word-picture identification in noise test, reported a much lower mean SRT (−11.55 dB SNR) when both the noise and speech were presented from the front (free-field). Although free-field testing was performed, the mean SRT was much better than that found in Study 5. The difference in the mean SRT found in the studies by Sheikh Rashid *et al.* (2017) and Yuen *et al.* (2019) and the mean SRT found in Study 5 is large. It is worth mentioning that Sheikh Rashid *et al.* (2017) and Yuen *et al.* (2019) used an adaptive procedure that estimates SRTs that corresponds to 50% correct responses, the adaptive procedure used in Study 5 estimates SRTs that correspond to 70.7% correct response. Besides, the differences could be drawn back to the difference in the tests' languages

(Dutch, Mandarin and Arabic) and other possible differences between children of different populations.

In general, it seems that the performance of children in SIN tests differs between languages. It also differs between different SIN tests that are performed in the same language. However, the mean SRT obtained from NH children using the PAAST is considered reasonable since it shows comparable findings at least to the MTT in noise, which the PAAST was inspired by in terms of the pairing of the words and the adaptive procedure used.

The availability of this information about the performance of NH Arabic-speaking children in a SIN test that uses an adaptive procedure is valuable for clinical and research purposes. It could also be used as a baseline for other Arabic SIN tests that are developed in the future.

6.5.3 Test-retest reliability of the PAAST

The reliability of the PAAST was assessed among the two groups of children (young and older children). The aim was for the children in both groups to perform the test for 3 rounds in 1 session in order to look for the test-retest reliability. A within-subject comparison was carried out to allow for defining any statistical differences between the performance of the children in different rounds by estimating the repeatability/critical value of the test. As mentioned previously, this information is important for the clinical practice where clinicians intend to test their patients under two different conditions. For instance, the critical value would help to find any statistically significant changes in the performance of a patient with and without a hearing aid. Additionally, it would support the clinician's decision to propose monoaural vs binaural hearing aids, monoaural cochlear implant vs bimodal fitting etc.

The test-retest reliability was assessed based on the findings of 2 rounds for the younger children (n=17) and 3 rounds for the older children (n=14) due to tiredness and/or unwillingness of some children in both groups to perform the second/third round. The reliability of the PAAST was compared to the reliability of the MTT. Lovett *et al.* (2012) assessed the test-retest reliability of the MTT in noise with 13 children aged 3.0-3.9 years (n=8) or 7.0-7.9 years (n=7). The children in their study performed the test in free-field over two sessions (once in each session). When both the speech and noise were presented from the front, the $SD\omega$ was 3.23 dB, indicating that the MTT in noise can predict an SRT that lies within ±6.3 dB of the child's true threshold. It also indicates repeatability of 8.8 dB for the MTT in noise. The correlation between the first and the second scores was low (r= 0.04). Summerfield *et al.* (1994) also assessed the reliability of the MTT in children but only in quiet conditions. The children performed the free-field testing twice on the same day, and the two test scores were compared to each other. A SDω of 2.5 dB (indicating repeatability of 6.8 dB) and a correlation of +0.95 between the two test scores were reported. Similarly, Lovett *et al.* (2013) assessed the reliability of the MTT in noise but with hearing-impaired children. When a pink noise was used, a SDω of 2.4 dB, a repeatability of 6.8 dB and a correlation of +0.73 were reported, which is similar to the reliability of the MTT in quiet (Summerfield *et al.*, 1994). The findings of Study 5 revealed that the PAAST showed better test-retest reliability than the MTT both in quiet and in noise with a SDω of 1.8 and 1.6, repeatability of 4.9 and 4.4 and correlation of 0.7 and 0.8 for the young and older children respectively. It is acknowledged here that the differences in the reliability measures between the PAAST and the MTT in noise as reported by Lovett *et al.* (2012) might be caused by the fact that the children in the study by Lovett *et al.* (2012) performed the test in two different sessions whereas in Study 5 the children were tested in the same session.

Comparable reliability was reported by Sheikh Rashid *et al.* (2017) for the online SIN screening test, which showed variability of 1.3 dB (repeatability of 3.6 dB). This repeatability differs by less than a dB from the repeatability obtained by the children of similar age groups in Study 5 (repeatability of 4.4 dB). On the other hand, Yuen *et al.* (2019) reported better reliability for the Mandarin spoken word-picture identification test in noise – adaptive (MAPID-A), which scored repeatability of 3.2 dB when children aged 4.8-5.2 years were tested (Yuen *et al.*, 2019). This is better than repeatability obtained from the young children (4.9 dB); however, the sample in Study 5 included younger children (aged 3 and 4 years) than children in Yuen *et al.* (2019); thus, the difference in age of the participating children in the two studies could be a cause for the variation in the reliability findings.

Other developed SIN tests for children used different methods to assess the reliability of the test, which make it difficult to compare the findings of Study 5 to those reported by them. For instance, Yuen *et al.* (2019) assessed the reliability by looking at the overlapping regions between the two 99% CI from the two rounds and reported the reliability in percentages. Denys *et al.* (2018) measured the SRT once from each child then divided the reversals that were used to calculate the SRT in two and ran mathematical calculations to estimate the reliability of the test.

In sum, the reliability measurements indicated that the PAAST yields to reliable outcomes when it is tested a few times on NH children. However, the findings of this study are only applicable to one condition, which is performing the test through headphones. It is recommended for a future study to assess the reliability of the PAAST when performed in free-field settings because it is expected that this condition is needed in clinical practice such as with children who refuses to put on headphones, children who use assistive hearing devices etc.

6.5.4 The normal range of SRTs

Due to the observed effect of increasing age on the SRT of NH children, preliminary normative data were provided for each age group separately. Some previous studies such as Sheikh Rashid *et al.* (2017), established age-correlated cut-off values per year age. In Study 5, the number of children per year age was small (Table 6.4); therefore, it was not convenient to estimate a correction factor per year age. As a result, the cut-off values were determined based on the determined normal ranges (−0.7 and −3.3 dB SNR for the young and older children, respectively).

One should be cautious when determining the cut-off value because different cut-off values would influence the accuracy of the test in differentiating between NH children and children with HL (Denys *et al.*, 2018). Denys *et al.* (2018) investigated different cut-off values for the digit triplet test in order to find the best cut-off value to differentiate between NH children and children with HL. They found that less strict cut-off values increased the diagnostic accuracy of the test and reduced the referral rate. However, it is not reasonable to consider less strict cut-off values unless they have been investigated for a particular test. Otherwise, children with HL might be missed. Therefore, investigating different cut-off values for each test and determining the optimal value that would lead to the best diagnostic accuracy as possible is recommended.

Due to the effect of age on children's SRTs, determining an age-specific cut-off value or a correction factor per year of age rather than a cut-off value for a wide age range is recommended. For example, Sheikh Rashid *et al.* (2017) found a significant improvement of 0.3 dB in the SRT per age year and based on this, they calculated cut-off values for each year of age. In the current study, although an improvement of 0.6 dB per year age was found, it was not possible to determine an age-specific cut-off value because of the small sample size. Therefore, at this stage, the previously proposed cut-off value was used for Study 6, which aims to investigate how well the PAAST can differentiate between NH children and children with SNHL. However, a future large-scale normative data study is recommended to be carried out to determine age-specific cut-off values for the PAAST.

6.6 Study 6 (children with SNHL): methods

Native Arabic-speaking children who were previously diagnosed with SNHL of different degrees (according to their audiograms) were recruited from the audiology departments of two main hospitals in Riyadh, KSA (The Security Forces Hospital and King Fahad Medical City). The inclusion criteria were: (1) children aged 3.0-18.0 years; (2) native Arabic speakers; (3) diagnosis of SNHL of any degree; and (4) acceptable language development that allows the child to respond to simple commands (picture pointing). The exclusion criteria were: (1) recent ear infection and (2) diagnosis of mental/cognitive disorders. Data collection continued for a period of 10 months (April 2018 to January 2019).

6.6.1 Sample size

The findings reported by Lovett *et al.* (2012) and Lovett *et al.* (2013) were used to build up the assumptions for the sample calculation of this study as follows. A mean SRT of around −4 dB SNR for young NH children was reported by Lovett *et al.* (2012) and a mean SRT of around +2 dB SNR for young children with SNHL was reported by Lovett *et al.* (2013). Based on this, a mean difference of around 6 dB is expected between SRTs in noise of NH children and children with SNHL. In addition, a SD of >4 dB was found for children with SNHL (Lovett *et al.*, 2013). However, Lovett *et al.* (2013) included children with moderate and higher degrees of SNHL whereas in Study 6, children with any degree of SNHL including mild degrees were recruited, so the SD was expected to be larger based on the expected variations in performance of children with different degrees of SNHL (mild to severe/profound). Therefore, the calculation was based on an expected SD of 8 dB for the children with SNHL. So, based on an expectation of a mean difference between the two groups (NH and SNHL) of 6 dB, a SD of 2 dB for the NH children and a SD of 8 dB for the children with SNHL, it was found that a minimum of 12 subjects with SNHL were needed to be tested in order to detect a mean difference of 6 dB with a power of 0.8 and alpha of 0.05 (Howell, 1997).

6.6.2 Procedure

Qualified audiologists working at the named hospitals recruited children who were attending the clinic for a routine follow-up appointment and obtained written consent from the parents and verbal consent from the children prior to any testing. Each child performed the PAAST once unaided monaurally. The SRT of the ear with better hearing sensitivity according to the PTA was measured with the PAAST. The test took place in a quiet room in the audiology department. The same equipment used to test NH children (Study 5) was calibrated and used in this study. The child was seated on a comfortable chair with the iPad placed on a table in front of him/her. In the beginning, the tester ensured that the child was familiar with all words by naming each picture and asking him/her to point to the named picture. This was to exclude the unfamiliar words before starting to test the child and to assess the familiarity of the words with children with SNHL. Afterwards, the child was informed that s/he was going to listen to sentences in the form of ("where is the (word)?") through the headphones and asked to touch the picture that represented the heard word. Children were asked to guess if they were not sure what the presented word was.

As mentioned in Section 6.2, the aim was to assess the ability of the PAAST to identify children with SNHL based on the distortion rather than the audibility component of HL. To do that, the loudness of the first presentation was set to the level where each child reported that the speech was comfortably loud (most comfortable level). Thus, the first presentation was audible to the children. By doing this, the audibility component of HL was reduced to some extent. In such a case, it is expected for the PAAST to identify the children based on the effect of distortion on their hearing abilities rather than the audibility.

Otherwise, if the initial loudness level was fixed at the same level for all children, it would be expected for the PAAST to identify hearing-impaired children based on both the audibility and distortion components. In this case, some children would be identified based only on the distortion effect (these are children with mild/moderate loss; because for those children, Class-A HL was not expected to interfere with the audibility of speech but be determined by the noise) whereas other children would be identified based on both the audibility and distortion components (those are children with more severe degrees of SNHL; because for those children, it was expected for the audibility of speech to be remarkably reduced and for the distortion component to be highly manifested). Identifying children who have Class-D only and children who have both Class-A and Class-D HL would make it harder to interpret the results in terms of the ability of the test to identify children who have Class-D HL. Besides, including Class-A HL in the identification would bias the results as it would enhance the performance of the PAAST in identifying hearing-impaired children and make the test look better than it is, given that it was motivated primarily as a test of distortion (Class-D HL) (sensitive to SNHL).

To ensure that the first presentation was audible to the child, the tester varied the loudness of the initial word using the sound mixer, which was connected to the iPad and the headphones, until the child reported that the word was comfortably loud to him/her. By doing this, it compensated for the direct contribution of the audibility component and emphasised the distortion component of HL, to some extent. Once the loudness of the initial presentation was set up for the child, the test was administered using the same adaptive procedure used for NH adults and children (see Chapter 5, Section 5.3.1). Copies of the audiograms of the participating children were provided to the author in order to look for a correlation between the PTA threshold and the SRT of the children, if any.

Ethics approval was provided by the ethics committee at the University of Southampton in the UK (identification number 40367) and from the participating hospitals.

6.7 Study 6 (children with SNHL): results

Although the sample size calculation suggested that including 12 subjects was enough to generate good statistical power for the purposes of the study, recruiting a larger number of children was intended to ensure including a variety of degrees of SNHL. However, over a period of 10 months, the testers managed to recruit only 17 children due to workload and/or difficulty in obtaining consent for participation. The SRT of one child, who had severe to profound SNHL (PTA threshold average of 3 frequencies was 90 dB HL), could not be obtained, probably because the child had difficulty identifying the correct word until the upper cap was hit, which caused the test to stop automatically. As a result, 16 native Arabic-speaking children (8 males and 8 females) aged 5 years and 8 months to 14 years and 3 months (mean= 9.0 years, SD= 2.5 years) who were known to have SNHL were included. For 8 children, the average PTA threshold was better in the right ear whereas for the other 8 children, the average PTA threshold was better in the left ear. Therefore, the SRT of the right ear was measured for 8 children whereas the SRT of the left ear was measured for the others.

All children had bilateral SNHL. The degree of SNHL varied among the children from mild to severe SNHL. According to the PTA thresholds (average of 0.5, 1 & 2 kHz), 6 children had mild SNHL, 5 had moderate SNHL, and the other 5 had severe SNHL (PTA thresholds for different frequencies and the SRTs obtained from the children are provided in Table 6.5). The aetiologies of the SNHL in the children were unknown to the author as the information was not recorded in the medical records of the children (as reported by the audiologists).

Table 6.5 The SRTs and PTA thresholds obtained from children with SNHL (n=16)

Subject	SRT (dB SNR)	PTA threshold (3 frequencies*) (dB HL)	PTA threshold at 0.5 kHz (dB HL)	PTA threshold at 1 kHz (dB HL)	PTA threshold at 2 kHz (dB HL)	PTA threshold at 4 kHz (dB HL)	PTA threshold (4 frequencies**) (dB HL)
1	-4.0	18	15	20	20	45	25
2	+7.0	65	45	65	85	85	70
3	-2.8	63	55	70	65	65	64
4	+18.0	75	60	75	90	85	78
5	+4.0	85	75	90	90	95	88
6	+13.7	73	70	75	75	75	74
7	-3.0	50	35	55	60	50	50
8	-2.8	55	35	70	60	60	56
9	+2.8	75	50	85	90	85	78
10	+6.3	47	40	55	45	50	48
11	-5.8	42	40	50	35	10	34
12	-5.0	12	25	15	-5	50	21
13	-3.3	28	25	30	30	25	28
14	+6.0	38	40	35	40	50	41

Subject	SRT (dB SNR)	PTA threshold (3 frequencies*) (dB HL)	PTA threshold at 0.5 kHz (dB HL)	PTA threshold at 1 kHz (dB HL)	PTA threshold at 2 kHz (dB HL)	PTA threshold at 4 kHz (dB HL)	PTA threshold (4 frequencies**) (dB HL)
15	+1.0	17	0	0	50	45	24
16	-4.3	47	35	40	65	55	49
17	Couldn't be tested	90	65	100	105	110	95

*Average threshold of three frequencies (0.5, 1 and 2 kHz)
**Average threshold of four frequencies (0.5, 1, 2, and 4 kHz)

6.7.1 The familiarity of the children with the words of the PAAST and the test duration

The 14 words of the PAAST were found to be familiar to all the participating children (100%). So, the 7 pairs of words were included in the test for all children. The mean duration required by the children to complete the test was 5.1 minutes (SD= 1.3 minutes). Because the age range of the children in the current study was almost similar to the age range of the older children (Group 2) in Study 5, the mean test duration of the two groups was compared statistically. The Mann-Whitney statistical test was used for that purpose because the two sets of data did not show normal distribution (Shapiro-Wilk test: p=0.04, p=0.008 for NH children and children with SNHL respectively). No statistically significant difference between the mean test duration of the two groups was revealed by the Mann-Whitney test (U= 150.5, p=0.8).

6.7.2 Mean SRT of children with SNHL

Sixteen children with different degrees of SNHL (mild to severe) completed one round with the PAAST. The SRTs obtained from the children ranged from −5.8 to +18.0 dB SNR with a mean of +1.7 dB SNR (SD= 7.0 dB). The mean SRT obtained from the children in the current study was compared to the mean SRT obtained from the NH children who were almost of the same age range in Study 5 (Group 2, older children). The SRTs obtained from the children with SNHL were compared to the SRTs obtained in R1 from Group 2 in Study 5 because the children in the current study performed the test only once. This way, the comparison would be appropriate as it looked at the performance of the children when they performed the PAAST for the first time.

The data (SRTs) obtained from the children with SNHL did not show normal distribution (Shapiro-Wilk test: p=0.03) whereas the data for the NH children were approximately normally distributed (Shapiro-Wilk test: p=0.4). Because one set of data was not normally distributed, the non-parametric statistical test (Mann-Whitney) was carried out. It revealed that the mean SRTs of the two groups were statistically significantly different (U= 4.0, p< 0.0001). However, since the independent sample t-test is considered a robust statistical test, it was also conducted. Conducting the independent sample t-test is useful to provide the 95% CI of the mean difference. The result of the t-test revealed that the mean SRTs of the two groups were statistically significantly different (t_{34}= -5.7, p< 0.0001) with a mean difference of −9.5 dB (95% CI of the difference: −12.9 to −6.1 dB). Figure 6.4 displays boxplots of the SRTs for NH children aged 6.0-12.0 years (R1) and the SRTs of children with SNHL and Figure 6.5 displays dot plots of the SRTs obtained from each NH child and each child with SNHL. It can be seen in the figures that the two groups of children are mostly distinguishable, with a small amount of overlap as some children with SNHL (n=5) provided SRTs that fell within the normal range (based on the previously determined cut-off value of −3.3 dB SNR). Two of these children had high-frequency SNHL (normal PTA thresholds up to 3 kHz), one had mild SNHL whereas the other two had moderate SNHL.

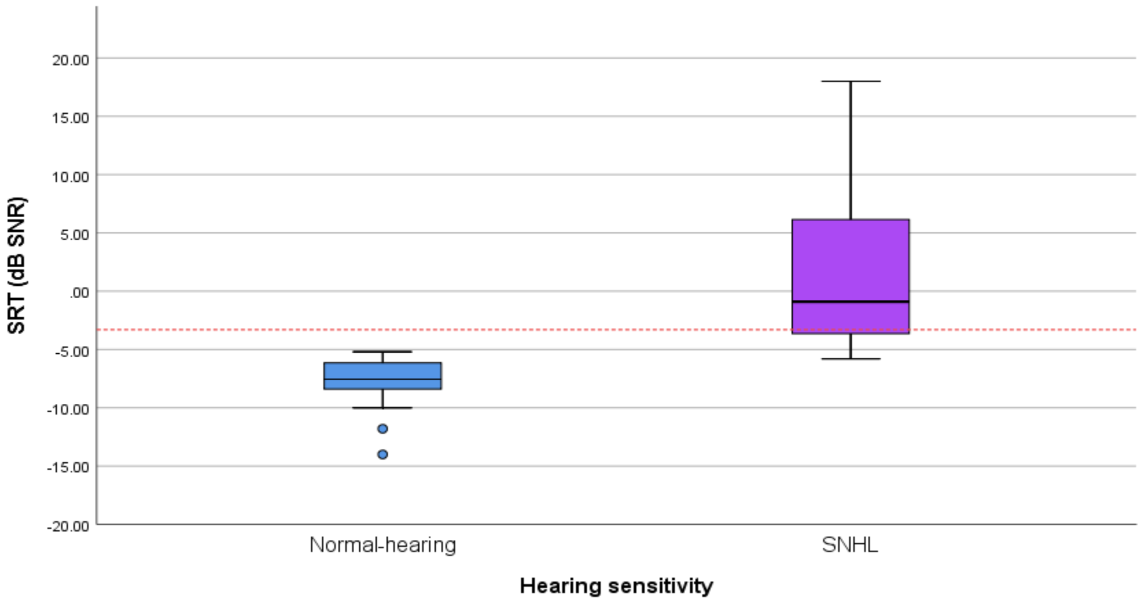


Figure 6.4 Boxplots representing the SRTs of NH children aged 6-12 years (R1) and the SRTs of children with SNHL. The dashed line at −3.3 dB shows the cut-off for NH (95% CI: −4.5 to −2.2 dB)

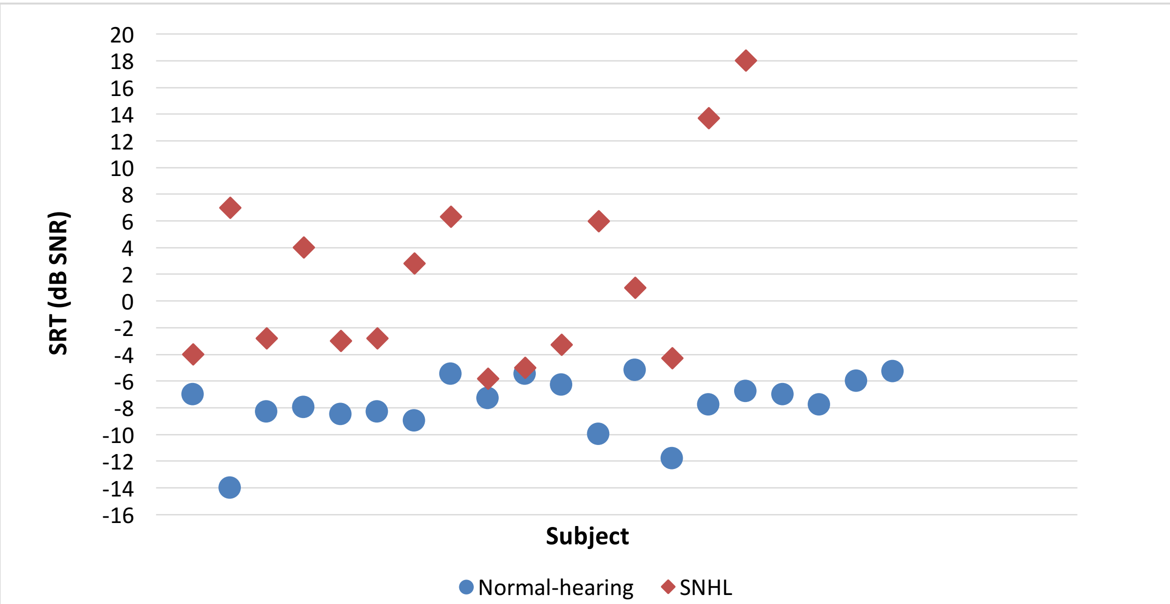


Figure 6.5 Dot plots showing the SRTs obtained from NH children (n=20) and children with SNHL (n=16). Each dot represents one child.

6.7.3 The relation between the SRTs and PTA thresholds

To assess whether there is a relationship between the SRTs and the PTA thresholds of the children with SNHL, Spearman's correlation coefficient was carried out. The non-parametric statistical test was used because the PTA data showed at least approximately normal distribution, whereas the SRT data did not show normal distribution (Table 6.6). Positive correlations were found between the SRTs and the PTA thresholds of different frequencies for children with SNHL, as shown in Table 6.7. The linear regression analysis revealed that the PTA thresholds of children with SNHL explain almost 30% of the SRTs obtained from those children (Table 6.7).

Table 6.6 Test of normality for the SRT data and the PTA data at different frequencies (Shapiro-Wilk test)

SRT	**0.5 kHz**	**1 kHz**	**2 kHz**	**4 kHz**	**Average (0.5, 1 & 2 kHz)**	**Average (0.5, 1, 2 & 4 kHz)**
p= 0.03	p= 0.9	p= 0.7	p= 0.4	p= 0.4	p= 0.6	p= 0.4

Table 6.7 Spearman's correlation coefficient and the coefficient of determination (R^2) of the SRTs and PTA thresholds of children with SNHL

PTA frequency	**0.5 kHz**	**1 kHz**	**2 kHz**	**4 kHz**	**Average (0.5, 1 & 2 kHz)**	**Average (0.5, 1, 2 & 4 kHz)**
SRT (rho)	0.63 (p= 0.008)	0.56 (p= 0.02)	0.63 (p= 0.008)	0.64 (p= 0.007)	0.61 (p= 0.01)	0.63 (p= 0.008)
SRT (R^2)	0.32	0.20	0.34	0.38	0.33	0.38

Scatter plots that represent the correlation between the SRTs and PTA thresholds of the different frequencies were almost similar to each other. An example is provided in Figure 6.6.

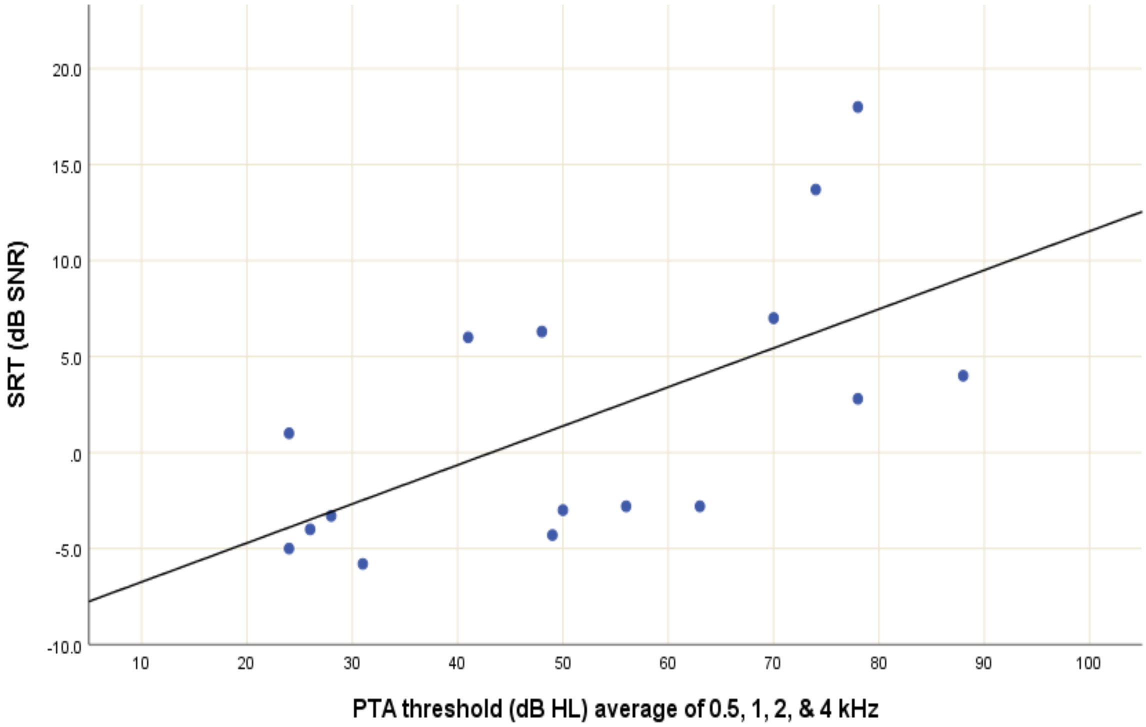


Figure 6.6 Scatter plot showing the correlation between the SRTs and PTA thresholds (average of 0.5, 1, 2 and 4 kHz) ($R^2 = 0.38$)

6.8 Study 6 (children with SNHL): discussion

In Study 5, preliminary normative data of the PAAST were provided for children and a cut-off value was determined to differentiate between NH children and children with HL. Afterwards, Study 6 was conducted to investigate the performance of children with SNHL when tested by the PAAST to find out how well the PAAST differentiates between NH children and children with SNHL and to look for a possible relationship between the SRTs obtained by the PAAST and the PTA thresholds of children with SNHL. The subsequent sections discuss the findings of Study 5 and 6 in terms of the familiarity of NH children and children with SNHL with the words of the PAAST and the duration required by the children to complete the test. Additionally, it discusses the findings of Study 6 in terms of the ability of the test to differentiate between NH children and children with SNHL and the relation between the SRTs of children with SNHL and their PTA thresholds.

6.8.1 The familiarity of the children with the words of the PAAST

The words used in the PAAST were found to be familiar to the NH children of all ages except for a few 3-year-old NH children (n= 3) where a few words (n= 5) seemed to be unfamiliar. However, the minimum number of pairs included in the children's testing was 5, and no common unfamiliar word was identified. For children with SNHL, all 14 words were familiar, so no pair was excluded from the testing, indicating that the words used are appropriate even for children with SNHL, who might have a limited vocabulary. However, since all children with SNHL in Study 6 were aged >5 years, it could be said that the words of the PAAST are expected to be within the vocabularies of hearing-impaired children who are older than 5 years. Although the words are expected to suit hearing-impaired children younger than 5 years, this information was not available from Study 6 and investigating it in a future study is recommended.

In general, it was concluded that there is no need to replace any of the words in the PAAST as they seem suitable to NH children in the KSA from the age of 3 years and hearing-impaired children from the age of 5 years. Besides, the familiarity of a non-Saudi child, who performed the PAAST with all the words, indicated that the PAAST could be used to test Arabic-speaking children in Arab countries other than the KSA.

The flexibility of the PAAST in terms of the ability to exclude pairs of words that are not familiar to a child is another reassuring point that there is no need for word replacement. It was reported that the MTT can be administered reliably with only 2 pairs of words as no effect was found for the number of pairs on the children's performance (Hall *et al.*, 2007; Lovett *et al.*, 2012). Therefore, it is expected that young children would be able to perform the PAAST reliably even if they were not familiar with some of the words.

6.8.2 Test duration

The short attention span of children requires keeping the task as short as possible to ensure reliable responses from the child. Elliott *et al.* (1979) reported that when adaptive procedure and proper reassurance were used, young children could stay attentive for the task and give reliable results with almost similar variability to older children and adults. This was seen with the PAAST, which uses an adaptive procedure and provides visual reinforcements, as young children provided reliable results that did not differ largely from older children. Besides, the duration required by young children to complete the PAAST (5.1 minutes) was very close to the duration required by older children and adults (4.6 minutes). Hence, it seems that children of different ages were engaged and attentive to the PAAST. Otherwise, if the children were distracted, the test would take them a longer time than adults to complete.

The findings of studies 5 and 6 revealed no difference between NH children and children with SNHL in terms of the duration required to complete the PAAST. It took approximately 5 minutes to obtain the SRTs from NH children and children with SNHL. This indicated that children with SNHL were able to complete the PAAST in clinical settings within almost the same duration required by NH children, who completed the test in almost controlled settings. A shorter duration was reported for the first and second versions of the semi-automated MTT, which was administered in quiet, (2.3 and 3.25 minutes respectively) (Ousey *et al.*, 1989; Palmer *et al.*, 1991). Similarly, some other paediatric auditory speech tests required a relatively shorter duration than the PAAST to be completed. For instance, children aged 9-12 years old completed the digit triplet in noise test in around 3.4 minutes (Denys *et al.*, 2018). Likewise, children aged 5-11 years old completed the online SIN screening test and the listening in spatialized noise-sentence test (LISN-S) test in around 3 and 3.11 minutes respectively (Cameron and Dillon, 2007; Sheikh Rashid *et al.*, 2017). The long duration of the PAAST is caused by the use of a child-friendly interface in which a pop-up screen of visual reinforcement appears after each response (explained in Chapter 5, Section 5.5.4).

On the other hand, other speech tests required a similar duration to be completed. For instance, 5-year-old children completed the Mandarin spoken word-picture identification in noise test (MAPID-A) in around 4.2 minutes (SD= 2.2 minutes), which shows a difference of less than a minute between the duration of the MAPID-A and duration of the PAAST (Yuen *et al.*, 2019). A similar finding was reported by Arnold *et al.* (2010) where children with cochlear implants required 5 minutes to complete the paediatric audiovisual speech test in noise, which uses the speech materials of the MTT. A duration of around 5 minutes is expected to be tolerable to children, particularly with the use of visual reinforcements and the use of technology (i.e. iPad), about which they are passionate nowadays. Additionally, the duration is considered reasonable for a test to be carried out in clinical settings. As a result, it is concluded that the testing duration of the PAAST is appropriate for use either for screening or diagnostic purposes in clinics.

6.8.3 SRTs of children with SNHL

For the purpose of Study 6, the SRTs of 16 children aged around 6-14 years who had different degrees of SNHL (mild to severe/profound) were measured. The better ear was tested once for each subject and revealed a mean SRT of +1.7 dB SNR (SD= 7.0 dB SNR). Studies that assess the SRT of hearing-impaired children usually measure the aided SRT (e.g. Rance *et al.*, 2007; Yang *et al.*, 2012; Caldwell and Nittrouer, 2013; Lovett *et al.*, 2013) whereas in Study 6, the unaided SRT was measured therefore the comparison with other studies might not be appropriate. This can be seen when the finding of Study 6 was compared to the finding of the study by Lovett *et al.* (2013). Although it was expected for the mean SRT children in Study 6 to be worse because the children performed the test unaided whereas in Lovett *et al.* (2013) the children performed the test with their hearing devices on, the mean SRT found in Study 6 was lower than the mean SRT obtained by children aged 3.0-6.3 years who had SNHL and performed the MTT in noise (mean= +3.14 dB) (Lovett *et al.*, 2013). This might be explained by the differences between the two studies in terms of methodology and participation criteria. In Study 6, the children's ages ranged from around 6-14 years, the children's degrees of HL ranged from mild to severe/profound, the stimuli were presented via headphones, and the unaided SRT was measured. On the other hand, in the study by Lovett *et al.* (2013), the children were younger, only children with moderate and worse degrees of SNHL were included and the aided SRT was measured in free-field. Hence, it is not surprising that the mean SRT obtained from children in Study 6 to be better than that obtained from the children in the study by Lovett *et al.* (2013), particularly with the inclusion of older children and children with mild degrees of SNHL in Study 6.

Wide variations were seen in the SRTs obtained from the children with SNHL (Figure 6.4). These variations can be explained by the variation in the degrees of their HL, which ranged from mild to severe/profound SNHL. Besides, the SRTs obtained from two children with severe/profound SNHL might have contributed a lot to this wide variation as they were highly positive (+13.7 and +18 dB SNR) in comparison to the SRTs obtained from the other children (Table 6.5). These highly positive SRTs could not be related directly to poor PTA thresholds because other children had similar audiograms and obtained better SRTs.

Possible explanations to these two highly-positive SRTs are as follows. First, there is a possibility that something went wrong in the way the test was administered. However, the possibility of this is very low, particularly as the testers are experienced audiologists and have administered the test to the rest of the subjects, who provided reasonable results. The second possible explanation is related to the procedure used in the testing, in terms of the loudness of the initial presentation, where the tester varied the loudness until the child indicated that the speech was clearly heard. It might be that these children had poor audibility of speech; thus they required the speech level to be increased to considerably high levels. In this case, the SRTs obtained might be their speech-in-quiet thresholds rather than SIN thresholds. This is possible because the test calculates positive SNRs by fixing the speech level and varying the noise level, which could result in reaching a level where the child no longer hears the noise. Third, it is possible that these two children had narrow dynamic ranges, thus with a small increase in the SNR, their SRTs improve greatly.

The SRT of +18 dB SNR, in particular, was questionable. For this case, it was difficult to tell whether this is an accurate calculation of the child's SRT or not because the upper cap for the SRT calculation is +20 dB SNR and the step size is 2 dB. Hence, there is a probability that the SNR hit the upper cap at some points but then decreased after two consecutive correct responses which might occur by chance. Although the adaptive procedure used reduced the probability that the SNR reduces by chance, and the number of choices is large (14), this still could occur in practice. If this were the case, the true SRT of that subject might be worse and the automated system of the PAAST is not expected to be able to calculate it because of the limited upper cap. This was seen with the child who was excluded from the study because no SRT could be obtained. However, it was not appropriate to exclude this finding (+18 dB SNR) because if the SRT obtained was true, the exclusion would bias the analysis.

Generally, the results showed that SNHL has a negative effect on the performance of Arabic-speaking children in understanding SIN. This can be seen even with children who had mild SNHL where they scored worse SRTs than NH children. However, the wide variation in the performance of the children with SNHL in SIN testing indicates that even with similar audiograms, SIN thresholds might vary mainly because of the extent of the distortion impairment. This finding supports the argument that PTA thresholds do not manifest the degree of difficulty in understanding SIN, and that SIN tests are better used for that purpose.

6.8.4 How well did the PAAST differentiate between NH children and children with SNHL?

The age range of the children with SNHL was similar to that of the older NH-children (Group 2 in Study 5), therefore the mean SRTs of both groups were compared. A statistically significant difference was found between the mean SRT of the two groups (mean difference= −9.5 dB). This finding indicates that the PAAST can differentiate between NH children and children with SNHL (Figure 6.4). However, some overlap was noticed between the two groups (Figure 6.5) where the SRTs of 5 children with SNHL fell within the proposed normal range (the proposed cut-off value= −3.3 dB).

Possible justifications to the overlap between the two groups are as follows. First, the procedure used in which the loudness of the initial presentation varied for each child to match a comfortable loudness of speech. However, it is acceptable to use this procedure because in clinical practice, when audiologists are interested in measuring the SRT of a patient for any reason, such as assessing the effectiveness of a hearing aid, they usually start the speech test at a comfortable loudness level and then continue with the test to estimate the threshold (Gelfand, 2009; Katz *et al.*, 2015). Besides, it was of interest to find the ability of the PAAST to detect children who had a distortion problem that interfered with their ability to understand SIN. For these reasons, the initial presentation level was varied for each subject. By doing this, the contribution of the audibility component was reduced, and the distortion component was emphasised, which made it harder to distinguish between the NH children and the children with SNHL. Otherwise, if the presentation level were fixed, the children who had low hearing sensitivity (worse PTA thresholds) would struggle to hear the words and thus would not be able to understand them. In this case, the test would distinguish between the two groups more easily, at least in principle. Nonetheless, even when using the harder condition (reducing the audibility component), the findings showed that the PAAST was still able, to some extent, to distinguish between the two groups. It is worth noting that for screening purposes, both the audibility and distortion components are important and should be taken into consideration. For that reason, when using the PAAST for screening purposes, the initial presentation level should be fixed (as it was when testing the NH children). In this case, it is most likely that the test will distinguish better between the two groups.

Second, the used cut-off value might be the reason for the overlap. The determined normal range is wide (−12.1 to −3.3 dB SNR), which increases the uncertainty about the optimal cut-off value that would lead to better results. As mentioned previously, testing of different cut-off values is recommended in order to determine the cut-off value that best differentiates between the two groups. Denys *et al.* (2018) stated that less strict cut-off values led to better diagnostic accuracy. The 95% CI of the lower limit of the normal range for Group 2 was −4.5 to −2.2 dB SNR. If −4.5 dB was used as the cut-off point, the SRTs of three of the children who seemed to have normal SRTs would be considered as abnormal SRTs. So, with this cut-off value, only two children with SNHL scored normal SRTs. These two children had normal PTA thresholds up to 3 kHz. This result is not surprising since it is known that the frequencies that are important for understanding speech sounds and that are usually associated with the SRTs are 0.5-2 kHz (Gelfand, 2009, 2010). However, by using this “less strict cut-off value,” the sensitivity of the PAAST would be reduced by increasing the false-negative cases and this would result in identifying children who have SNHL as having normal hearing. For that reason, it was preferred not to change the proposed cut-off value (−3.3 dB), for now, and to recommend a future study to be conducted with a larger sample size of NH children to provide more accurate normative data. The inclusion of a larger number of children would narrow the normal range and increase the certainty of the appropriateness of the cut-off value. In addition, assessing different cut-off values to determine the optimal cut-off value that could differentiate between NH children and children with SNHL with high sensitivity and specificity is recommended.

Third, it is possible that the children who obtained SRTs that are within the normal range had relatively better frequency selectivity, temporal resolution or cochlear compression functions than children who had poor SRTs. It was noted in Chapter 2, Section 2.2.1, that reductions in such functions are the main causes for the decreased ability to process speech stimuli in individuals with SNHL. Hence, if those functions were not severely affected, the individual performance in speech discrimination is not expected to be reduced largely. On the other hand, it is possible that these children used some coping strategies that would help them to overcome the challenges they faced in understanding SIN. As mentioned earlier, the frequency of word usage could influence the child's ability to understand in challenging listening conditions. It is possible that those children found the task simple and were able to fill in the gaps, particularly if they were aged 6+ years. Although each pair shared the same vowel, which was supposed to make the identification of the word harder if the consonants were not identified correctly, a child can identify either the initial or the last consonants and the vowel and use this information to fill in the gap and identify the presented word correctly. This raises an important point worth investigating in a future study, which is the self-reported difficulties understanding SIN. This would help in exploring possible usage of coping strategies, thus understanding the reason behind the unexpected good SRTs that were obtained by children who had SNHL.

Fourth, in the current study, speech-shaped noise was used as the masker. It is possible that other types of noise such as babble noise might distinguish better between NH children and children with SNHL. This is worth investigation in a future study to find the best type of noise to be used with PAAST that would lead to better results.

In general, the finding of this study is promising as it showed that the PAAST could differentiate between the two groups of children (NH and hearing-impaired) (Figures 6.4 and 6.5). This indicates that the PAAST could potentially be used for hearing screening purposes. The noticed overlap between the SRTs obtained from children with NH and children with SNHL is not expected to prevent its usage for hearing screening purposes because the area of the overlap is small (less than 3 dB) as can be seen in Figure 6.5. As discussed above, the specified cut-off value might cause this overlap because it was determined based on findings obtained from a small sample of NH children. Hence, altering the cut-off value could possibly increase the sensitivity of the PAAST to SNHL by finding the optimal cut-off value that best discriminate between the two groups of children. It is important to note here that the sensitivity and specificity of the PAAST was not assessed in this study. However, even if it was found that the PAAST is not sensitive to mild cases of SNHL, it could be used to screen for more severe cases as it showed potential to detect those children. Besides, the PAAST could be used in a hearing screening protocol together with other hearing screening tests such as PTA or otoacoustic emission. The inclusion of the PAAST in a hearing screening programme is of a high value as it shows the children's ability to function in noisy listening environments which is common in the daily life.

At this stage, the diagnostic significance of an individual score that occurs at or near the specified cut-off value needs to be interpreted with caution, taking into account the overlap between the two groups of children. For clinical practice, audiologists are recommended to repeat the test in such cases and support the findings with other audiological tests.

6.8.5 The relation between SRTs and PTA thresholds

Previous studies showed a highly positive correlation between PTA thresholds and SRTs in quiet. For instance, Palmer *et al.* (1991), Summerfield *et al.* (1994) and Hall *et al.* (2007) reported a correlation coefficient (r) between the SRT obtained from the MTT in quiet and PTA thresholds that range between 0.7 to 0.9. However, the correlation between PTA thresholds and SRTs in noise particularly for individuals with SNHL varied in literature from low (e.g. Festen and Plomp, 1983; Divenyi and Haupt, 1997), moderate (e.g. Glasberg and Moore, 1988; Kramer *et al.*, 1996; Beattie *et al.*, 1997; Denys *et al.*, 2018) to high correlations (e.g. Smoorenburg, 1992; Hagerman, 2002; Smits *et al.*, 2004). This is because the ability to detect pure tones is mainly affected by the audibility of the sounds whereas understanding speech sounds, particularly in the presence of noise, is affected by both the audibility and the distortion of the sounds that result from the SNHL as discussed in Chapter 2, Section 2.2.1 (e.g. Plomp, 1978; Gelfand, 2010). Therefore, observing a strong linear relation between SRTs in noise and PTA thresholds is not expected.

The findings of Study 6 revealed a moderate correlation between the PTA thresholds of different frequencies and the SRTs of the subjects, where the PTA thresholds explained around 30% of the SRTs of the children. Although the method used in Study 6 attempted to reduce the contribution of the audibility component of SNHL on the obtained SRTs, the contribution was not fully eliminated. Thus, the audibility component is expected to have at least a small effect on the SRTs and thus have a contribution to the correlation between the PTA thresholds and the SRTs. Besides, it is expected that children who have worse degrees of SNHL (worse PTA thresholds) would have more distortion effect, thus would have poorer SRTs than children with mild/moderate degrees of SNHL. This might explain the correlation found in the current study.

The findings suggest that the PTA thresholds could not fully explain the SRTs, particularly as some children with similar PTA thresholds obtained SRTs that were largely different from each other. This highlights the need for SIN tests to assess the ability to understand SIN. To demonstrate the value of SIN testing, two examples from the data of Study 6 are discussed as follows. First, Subject 14 and Subject 16 (Table 6.5) had almost similar PTA thresholds at the frequency range 0.5-4 kHz (moderate SNHL with 5 dB differences between the thresholds of the two children at each frequency except for the 2 kHz where Subject 16 had a threshold that is 25 dB worse than that of Subject 14). If the PTA thresholds could predict the SRTs, it would be expected for the two children to obtain similar SRTs or at least, it would be expected for Subject 16 to score a poorer SRT as the child had a worse threshold at 2 kHz than Subject 14. However, the findings showed that Subject 16 obtained a much better SRT (−4.3 dB SNR) than Subject 14 (+6.0 dB SNR) despite similar audiograms. Additionally, the SRT obtained by Subject 16 fell within the normal range (as determined in Study 5), although the child had moderate SNHL (according to the PTA thresholds). This indicates that Subject 16 might not have experienced serious difficulty in understanding SIN regardless of the degree of HL.

Subject 4 and Subject 9 showed another example that similar audiograms do not necessarily indicate that children would have similar SRTs in noise. Both children had similar PTA thresholds (severe SNHL with 10 dB HL differences at 0.5 and 1 kHz and no difference at 2 and 4 kHz), thus it is expected for the two children to perform similarly in the SIN test, particularly as they had the same PTA thresholds at high frequencies. However, Subject 9 scored a much better SRT (+2.8 dB SNR) than Subject 4 (+18 dB SNR). The findings illustrate that PTA thresholds are not accurate predictors of speech understanding in noise and do not reflect communication difficulties experienced in noise. SRTs in noise are probably a better indication of the difficulties experienced in everyday life situations than PTA thresholds. Therefore, to understand the child's ability to communicate in noisy situations, directly measuring this ability using SIN testing is recommended.

In sum, it could be concluded that the PAAST is a fast and reliable SIN test for children. It can be used clinically to measure SRTs for diagnostic and rehabilitative purposes. Additionally, it can be used for hearing-screening purposes once the accuracy measurements of the test in terms of the sensitivity, specificity, and positive and negative predictive values are assessed.

6.9 Conclusions

The current chapter included two studies. Study 5 assessed the performance of NH Arabic-speaking children in SIN testing using the PAAST, whereas Study 6 assessed the performance of Arabic-speaking children who had SNHL. The findings of the two studies are summarised as follows:

- All words used in the PAAST were familiar to 93% of the NH children aged 3-12 years and 100% of the children with SNHL aged 6-14 years
- NH children from the age of 3 years and children with SNHL could complete one run of the PAAST in a short time (around 5 minutes)
- Reliable SRTs can be obtained from NH children as young as 3 years old
- An effect of age was noticed on the children's performance in SIN testing
- Preliminary normative data and cut-off values were provided for children of two age groups (3-6 and 6+ to 12 years old)
- The PAAST was able, to some extent, to differentiate between NH children and children with SNHL
- A moderate correlation was found between PTA thresholds of different frequencies and SRTs in noise when measured using the PAAST.

Chapter 7 Study 7: The usability of the PAAST and the feasibility of using the PAAST in a preschool hearing screening in schools in the KSA

7.1 Overview of the chapter

The Paediatric Arabic Auditory Speech Test (PAAST) was implemented in an iPad application to ease the accessibility to the test and allow for its usage for childhood hearing screening purposes. This chapter outlines the exploration of the usability of the PAAST iPad application when operated by parents and teachers. It also explores the feasibility of using the PAAST to screen the hearing of kindergarten children at schools in the Kingdom of Saudi Arabia (KSA).

7.2 Introduction

Chapter 3 concludes that the age of identification (AOI) with sensorineural hearing loss (SNHL) in children in the KSA is high, with some children identified after school age. The main possible reasons behind that are late-onset SNHL and/or acquired SNHL, difficulty accessing audiology services and limited awareness of SNHL and the consequences of untreated or lately addressed SNHL in children. Based on these findings, it was concluded that the universal neonatal hearing screening (UNHS), which was launched recently in the country, might be inadequate to identify a high percentage of cases of SNHL, therefore options for hearing screening later in childhood were recommended. Two options were investigated in this research: (1) to provide a hearing screening tool that could be easily accessed and used by non-audiologists to screen the children's hearing immediately in case of suspicion without the need for waiting a long time to be referred and seen by an audiologist; (2) to apply a targeted hearing screening to kindergarten children who are at risk of SNHL.

It was noted before (Chapter 2, Section 2.3) that the difficulty in perceiving speech in the presence of background sounds is the most challenging for individuals with SNHL (Davis, 1989; Kramer *et al.*, 1996) and that speech-in-noise (SIN) tests have been recommended to be used to assess this difficulty. Hence, it seems convenient to employ such tests to screen for SNHL. The PAAST was therefore developed carefully to match the requirements of screening tools in order to propose its usage for children's hearing screening. For instance, the PAAST was implemented in an iPad application for ease of accessibility to the test without the need for special equipment. Additionally, the procedure of the PAAST is simple, i.e. the hearing threshold is measured automatically without the need for interference from the person administering the test. Moreover, the test is expected to be easily administered by individuals who have no experience of audiology services such as parents and teachers. Furthermore, the test could be completed in a short duration.

The PAAST could be used any time by parents, teachers and/or general practitioners in case of suspicion about the child's hearing abilities. It could also be used for a broader and more formal hearing screening programmes such as a preschool hearing screening programme. Such programmes are important to identify children with SNHL who are missed by the UNHS programme, who develop delayed onset SNHL, who acquire SNHL and/or who have a mild degree of SNHL, which is not targeted by the UNHS targets. As mentioned previously in Chapter 2, Section 2.2.5, school-entry hearing screening programmes (SEHS) were implemented successfully in high-income countries (HICs) such as the UK and the USA (Bamford *et al.*, 2007; Fortnum *et al.*, 2016) and are seem to be important in countries in the Middle East particularly because of the high practice of consanguinity which is a well-known cause for late-onset SNHL (Zakzouk *et al.*, 1995; Schrijver, 2004; Shawky *et al.*, 2013). It was also mentioned in Chapter 2, Section 2.2.6 that a school screening program has been launched recently in the KSA (Ministry of Health, 2019d), and that it is noticed that the time gap between the UNHS and the school hearing screening is large (6 years), which would result in delaying the identification of children who have SNHL. Therefore, it was suggested to apply a targeted hearing screening to children who are at risk of SNHL at kindergarten age (3-5 years) at schools.

Targeted screening programmes are preferred when it is easier and cheaper to identify people at risk (Del Mar *et al.*, 2008). It is accepted and applied to screen for some diseases in HICs. For example, in the UK, diabetic eye-screening targets only diabetic people as they are of a high risk of developing eye diseases and blindness. The screening programme offers diabetic people from the age of 12 years in the UK an annual eye-screening test. The programme is considered successful as diabetes is no longer the leading cause of blindness in the UK (National Health Service, 2019). Hence, it seems that applying targeted screening programmes is beneficial to avoid or reduce the negative impact of diseases and disorders as much as possible. It is therefore expected that applying a targeted hearing screening programme would help in identifying children with SNHL as early as possible. The feasibility of applying such a screening is explored in this chapter (Study 7).

It was important to inspect the feasibility and usability of the PAAST interface to determine whether individuals who have no experience of audiology services could use the PAAST to carry out simple hearing screening tests on children. The following section describes the feasibility and usability testing of this tool.

7.2.1 Usability and feasibility testing

Technical systems such as websites, software and applications need to be efficient, intuitive and easily usable by end-users (Adler and Winograd, 1992; Nielsen, 1993; Preece, 1993; Brooke, 1996). It was necessary to ensure at this early stage that the PAAST interface is user-friendly and that the test can be administered easily by individuals who have no experience with audiology services and might have little experience with technology because it is aimed in the future to release the PAAST iPad application for public use. If the PAAST will be used in future for screening programmes, the last thing that is wanted to be found at that time is that the design of the PAAST application is complex and the procedure of administering the test is difficult. Although the PAAST interface was designed carefully to ease its administration, it is still difficult to predict how other people are going to find it. For that reason, the usability of the PAAST was explored in Study 7.

Usability is related to the effectiveness and efficiency of the system and the satisfaction of the users. Usability testing is a procedure where data is collected about the users' attitude towards a system/interface in terms of ease of use, learnability, the efficiency of use and willingness to use (Adler and Winograd, 1992; Nielsen, 1993; Preece, 1993; Brooke, 1996). One method of assessing the usability is through survey evaluation, which aims to derive information about the users' subjective opinions of an interface after they use the system to perform some tasks. This could be done through questionnaires or interviews. By conducting the usability testing, any identified problem/difficulty could be addressed to improve the system of interest. Therefore, it has been recommended for the usability testing to be conducted at early stages, particularly in the beta-version, where the development is completed but the system/application is not yet released for public use. Conducting usability testing after making any amendments to assess the usefulness of the changes is also recommended (Adler and Winograd, 1992; Nielsen, 1993; Preece, 1993; Brooke, 1996).

Several reliable questionnaires are available for usability testing such as the after-scenario questionnaire, the poststudy system usability questionnaire, the usefulness, satisfaction and ease of use, the computer system usability questionnaire etc. (Lewis, 1995; Lund, 2001; Lewis, 2002). A commonly used questionnaire is the system usability scale (SUS) questionnaire, which was developed by Brooke (1996). The SUS is used to assess the usability of a variety of websites, software, medical technologies etc. (e.g. Maidment and Ferguson, 2018) because of the following reasons. First, it showed high validity even for small samples (Tullis and Stetson, 2004; Bangor *et al.*, 2008). Second, it is known for its flexibility in assessing a wide range of interface technologies. Third, it is quick, short and easy. Fourth, it is easily scored and provides a single understandable score on the scale that quantifies the users' interaction with a given system (Bangor *et al.*, 2008).

In addition to the usability testing, it was also important to investigate the feasibility of applying a targeted hearing screening at schools in the KSA prior to proposing using the PAAST for that purpose. Conducting a feasibility study, which is a small study that is conducted prior to the main study, allows for investigating whether the main study can be done (National Institute for Health Research, 2012). In other words, it ensures that the study can be implemented practically. Findings of feasibility studies help researchers to identify problems, overcome them and build a foundation for their main study (Arain *et al.*, 2010; Tickle-Degnen, 2013). Therefore, investigating the feasibility of using the PAAST to screen the hearing of kindergarten children by teachers at schools in the KSA was considered in this PhD project. It is expected that the findings of this study will provide useful information about the practicality of applying a targeted hearing screening at schools by showing the acceptability of school staff members to such a practice, by showing the acceptability of children to having their hearing screened at schools, and by raising any possible difficulties and obstacles. Knowing this information at this stage would help in overcoming the difficulties and finding solutions to come up with the best plan that could be administered easily and practically.

7.2.2 Aims

The current study aims to:

1. investigate the **usability** of the PAAST when used by parents of children aged 3-6 years and teachers of kindergarten children (aged 3-6 years)
2. investigate the **feasibility** of applying a targeted hearing screening to kindergarten children using the PAAST in the KSA among teachers.

For the **feasibility part**, it was aimed to investigate the following:

- Acceptability of headteachers to the application of children's hearing screening at schools
- Willingness of teachers to administer the hearing screening
- Willingness of children to perform the hearing screening at schools
- Time required to screen the hearing of one child
- Overall time required by each teacher to complete the screening.

7.3 Methods

The current study consists of three parts: (1) the usability of the PAAST when used by parents, (2) the usability of the PAAST when used by teachers, and (3) the feasibility of screening the hearing of kindergarten children by teachers using the PAAST at schools in the KSA.

7.3.1 The usability of the PAAST when used by parents (online questionnaire)

In this study, Saudi parents who have at least one child aged 3-6 years with Arabic as their native language and who have accessibility to an iPad were recruited through social media (Twitter and Facebook). The study adverts were available on social media for two months. The study adverts included a link through which the subjects declared their agreement for involvement in the study (via Survey Monkey) and provided their emails for the researcher to email them back with the information sheet. Through emails, the researcher provided the subjects with the information sheet, an invitation to the application through TestFlight-app[6], an instruction sheet that describes the procedure of administering the test in detail (Appendix G) and a link to the usability questionnaire to be filled in after using the PAAST (via Survey Monkey) (the questionnaire is provided in Appendix H). Although 12-14 participants are enough to get reasonably reliable results with the SUS questionnaire (Tullis and Stetson, 2004), a minimum of 20 subjects is needed to discover a minimum of 95% of possible problems/difficulties with the system (Faulkner, 2003). Therefore, the aim was to include a minimum of 20 subjects.

For the reasons mentioned earlier in Section 7.2, the SUS questionnaire was used in Study 7 as the outcome measure for the evaluation of usability. The SUS consists of 10 items that look for the efficiency, effectiveness and satisfaction of the system under investigation. The answers are given in a five-point Likert scale from “strongly disagree” to “strongly agree”. Scores of each item range from 0-4, where the scores for items 1, 3, 5, 7 and 9 are calculated as the scale position minus 1 whereas the scores for items 2, 4, 6, 8 and 10 are calculated by subtracting the scale position from 5. A composite score, ranging from 0-100, is obtained by multiplying the sum of all item scores by 2.5. Scores ≥68 are considered above average and scored less than 68 are considered below average (Sauro, 2011).

In addition to the SUS items, subjects were asked to report their positive/negative comments and explain the difficulties they found, if any, as free text. Moreover, subjects were requested to provide their contact information if they were happy for the researcher to contact them by phone or mail

---

[6] TestFlight is a beta testing application that is used to invite people to test a beta version of developed applications before releasing them for public use.

to further discuss their opinion about the PAAST application. It is worth mentioning that the items of the questionnaire were translated into Arabic for the subjects to be able to understand it. The translation was not expected to affect the reliability of the questionnaire as it was reported that the SUS has been translated to other languages such as Spanish, French, and Dutch and showed similar reliability to the English version (Brooke, 2013).

Subjects who agreed to participate downloaded the PAAST application on their iPads, followed the instructions to administer the test on their children and then completed the SUS questionnaire. It was recommended for the children to use iPhone earphones because they were checked in the lab for the maximum output level.

Ethics approval was provided by the ethics committee at the University of Southampton in the UK (identification number 45826).

7.3.2 The usability of the PAAST when used by teachers and the feasibility of applying the hearing screening using the PAAST at schools in the KSA

For this part of the study, 5 private and 3 public schools, who have kindergarten children aged 3-6 years, in Riyadh, KSA were approached. Two of the public schools refused to participate whereas the third provided verbal approval to participate but failed to provide a written letter; thus, it was excluded. One of the private schools was quite busy at the time of the data collection, which made it difficult to contact the teachers; thus, the school was excluded. As a result, teachers of kindergarten children were recruited from 4 private schools located in different areas of Riyadh, KSA (central, north, east, and west).

A sample size between 24 and 50 is recommended for feasibility studies (Julious, 2005; Sim and Lewis, 2012); therefore, the aim was to recruit a minimum of 24 teachers. It is worth noting here that, in the KSA, kindergarten children are only taught by female teachers; therefore, it was expected for the sample to include only female teachers. The researcher (the author) visited the schools, recruited teachers and provided them with information sheets. Teachers who were happy to participate were asked to sign a consent form and were provided with instruction sheets to guide them in using the PAAST application to test children in their classes. Teachers were also provided with some documents to be sent to carers of all children who are in their classes. These documents were an information sheet, which clearly outlines the child's role within the study, a consent form, and a checklist of risk factors of SNHL. The risk-factor checklist included the 5 risk factors of SNHL that were found in Studies 1 and 2 (Appendix B). Parents/carers who were happy for their child to take part were asked to sign a consent form and to answer the checklist. It was stated clearly in the information sheet that no results could be provided about the child's hearing because the test is still under development.

The developed risk-factor checklist was used as a first-line screening tool to identify children who are at risk of SNHL. Teachers were requested to use the PAAST to test only children who appear to have at least one of the risk factors (as reported by carers) and whom parents were happy for them to be tested. However, teachers were also informed that if no child in their classrooms appeared to have any of the risk factors, they could then perform the test on children who had consent to be tested and were free of risk factors. This was to avoid losing a subject (teacher) who gave the consent to participate since the ultimate goal was to explore the feasibility of performing a hearing screening using the PAAST at schools.

An iPad with the PAAST installed was provided to the teachers together with an instruction sheet and the SUS questionnaire. Each teacher was asked to read the instruction sheet and follow it carefully to administer the test on the children. Before starting the test, teachers ensured that the children were happy to be tested. Children who refused the testing were not tested despite their carers' consent. The testing was done in a quiet room at each school. Some teachers preferred to seat all the children who had consents together in the testing room and tested them one by one whereas other teachers called them child by child from their classrooms. In cases where all the children were seated in the test room, children who were waiting their turns were kept busy with a colouring task to keep them quiet. After completing the testing, each teacher filled out the SUS questionnaire (Appendix I). In addition to the SUS items, there were questions about the number of tested children, the test duration for each child, the duration needed to test all the children and an open question for the teacher to report her positive and/or negative thoughts about the PAAST application. Answers to the SUS items allowed for exploration of the usability of the PAAST iPad application whereas answers to the extra questions allowed for exploring the feasibility of performing the test at schools in the KSA.

Ethics approval was provided by the ethics committee at the University of Southampton in the UK (identification number 45826) and from the participating schools.

7.4 Results

7.4.1 Usability analysis (parents – online questionnaire)

Sixty-four subjects responded to the study advert; 2 of them did not consent to participate. Of the 62 who consented to participate, 18 did not provide their email addresses, so there was no way to further contact them to provide them with the test application. Another 18 did not respond after they received the invitation to the test application regardless of receiving 2 reminder emails. As a result, 26 parents (4 males and 22 females) aged 24-58 years (mean= 36.5 years, SD= 8.4 years) were included.

Twenty-six parents had administered the PAAST to their children and filled out the SUS questionnaire. The overall usability score was 81.3 (Figure 7.1). The majority of the parents found the application easy to use (n=25, 96%), felt very confident using it (n=25, 96%), imagined that most people would learn to use it very quickly (n=25, 96%), found the various functions of the applications well integrated (n=22, 85%) and thought that they would like to use it if needed (n=25, 96%). On the other hand, 4 (15%) parents found the application unnecessarily complex, 3 (12%) thought that they would need assistance to be able to use it and 2 (8%) reported that they needed to learn a lot of things before they could get started with the application. The mean duration needed to complete the test by an individual child was 4.8 minutes (SD=1.2 minutes).

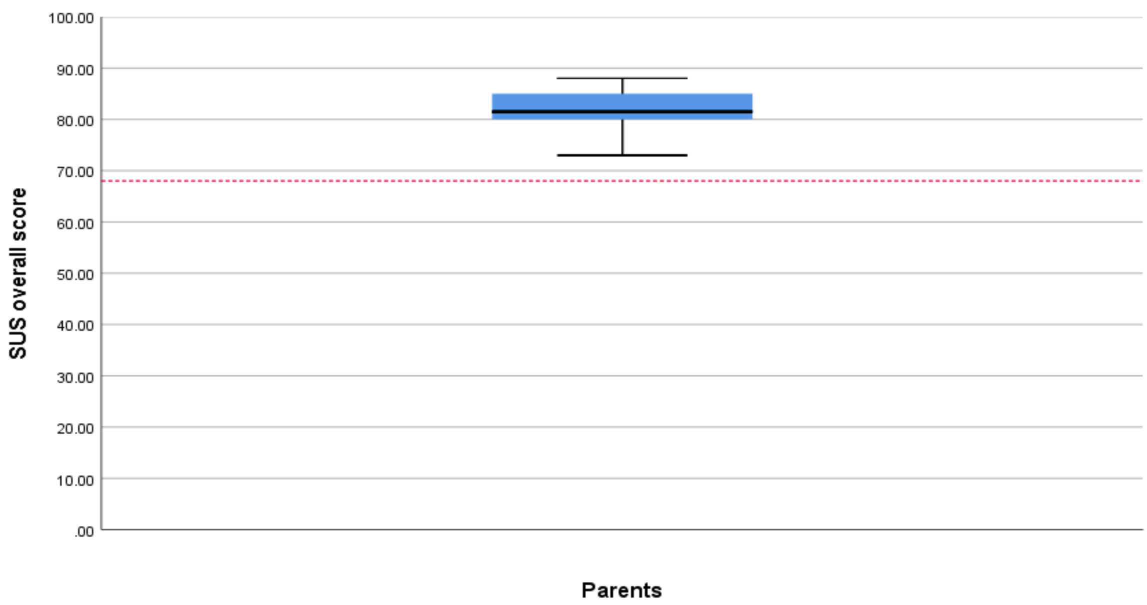


Figure 7.1 Boxplot showing the overall SUS score of parents. The dashed line denotes a score of ≥68, which is considered above average (Sauro, 2011).

No free text comments were provided. Therefore, 2 subjects were contacted by phone and were asked about their positive and negative feedback about the PAAST and 3 positive points were raised by them. First, they found the implementation of the PAAST in an iPad application a good idea because their children looked at it as a digital game rather than a hearing test and thus, they were happy to perform it without any effort from the parents. According to them, this made it a good option in case of suspicion about the child's hearing as a screening test at home. They also thought that it would encourage children to visit the audiology clinic, if needed, as the PAAST gave them a good impression about hearing assessments. Second, the use of the visual reinforcement (happy face) helped in attracting the children to complete the test as they noticed their children smiling and looking at the parents proudly when the happy face popped up. Third, in their opinion, the ability to exclude unfamiliar words would allow children to perform the test more confidently; otherwise, the task might be challenging to the child.

On the other hand, some negative comments were provided. First, users would prefer a detailed step-by-step instruction to be available within the application. Second, they reported that having 14 pictures to choose from might be confusing to the children, so they recommended reducing the number of choices. Third, they suggested having an "I don't know" sign/picture for the child to use in case s/he was not sure what the presented word was rather than asking the child to guess. According to them, guessing the word might cause random rather than accurate selections of the heard words. Besides, it might increase the test duration as the child needs time to guess. Fourth, they suggested reducing the number of times the visual reinforcement screen appears (maybe show it after 3-4 responses to reduce the testing duration). Finally, they preferred to have an option to pause the test in case the child lost attention, wanted to use the toilet etc.

7.4.2 Usability analysis (teachers)

Twenty-four female teachers out of 43 (56%) consented to participate in the current study, 9 from the first school, 3 from the second and 6 from each of the third and fourth schools. Thus, in total, 24 female teachers aged 25-37 years (mean= 29.8 years, SD= 3.3 years) were included. One of the participating schools is only for kindergarten children. It is relatively small with a few numbers of teachers and children. It has one classroom for kindergarten-1 (KG1) (children aged 3-4 years), one classroom for KG2 (children aged 4-5 years) and one classroom for KG3 (children aged 5-6 years). The other 3 schools were relatively larger than the first school. They accept children from kindergarten age to year 12. It is common in the KSA to assign 2 teachers for each kindergarten classroom (a main teacher and an assistant teacher). This was almost seen in all the 4 participating schools.

Twenty-four questionnaires were available for usability-testing analysis. The PAAST scored 81.5 (Figure 7.2) on the SUS. The majority of the teachers found the PAAST easy to use (96%), felt very confident using it (96%), imagined that most people would learn to use it very quickly (92%), thought that they would like to use it if needed (92%) and found the various functions of the applications well integrated (80%). Only 1 (4%) subject thought that there was inconsistency in the test or needed a lot of learning before being able to use it. Conversely, no teacher found it complex or needed assistance to use it. No free text comments were provided. However, the researcher received positive verbal feedback from several teachers who were keen to have it released as soon as possible for public use. For example, a teacher said, "What an impressive application! I will definitely use it with my students." Another said, "I hope to have it soon in our school." A further 7 teachers had almost the same comment: "I used to have some students that were not responsive, and I used to be confused whether they were intentionally ignoring my instructions, or if they had hearing difficulties. The availability of such an application would help in providing me with an objective evidence about their hearing abilities. And in case it shows that the child has a hearing problem, I can confidently raise the issue with the parents." One headteacher said: "Try your best to release it with no delay; there is a big need for such a hearing test in the field."

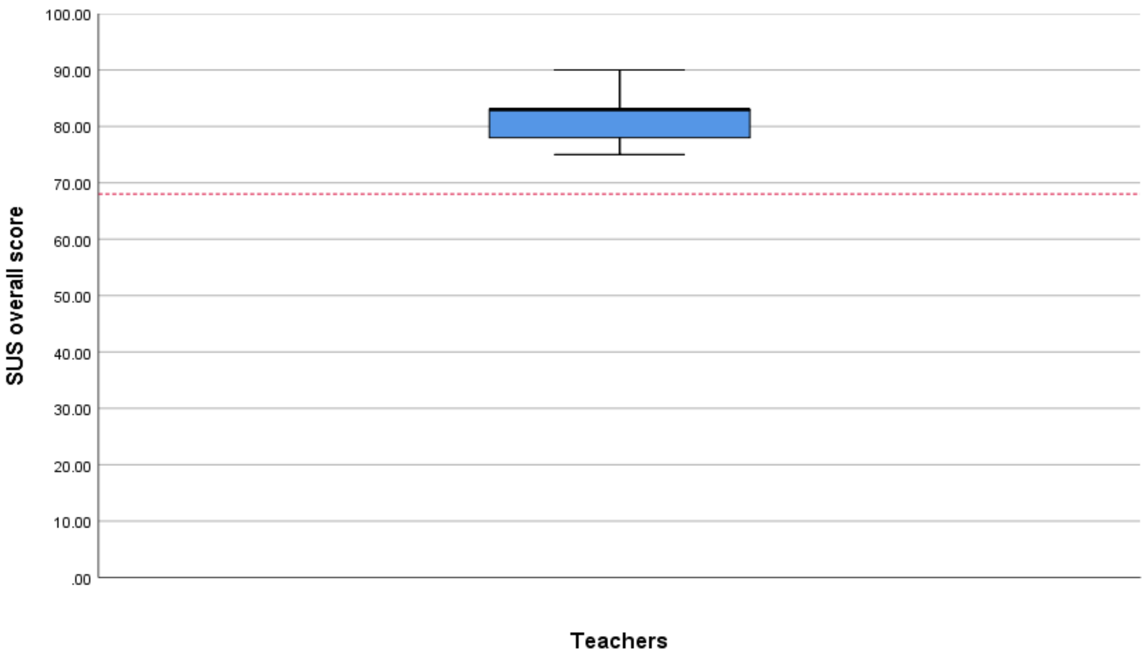


Figure 7.2 Boxplot showing the overall SUS score of teachers. The dashed line denotes a score of ≥68, which is considered above average (Sauro, 2011).

7.4.3 Feasibility of administering a hearing screening using the PAAST in schools in the KSA

Twenty-four teachers out of 43 (the total number of teachers of kindergarten children at the 4 schools) participated in the current study indicating a response rate of 56%. The total number of children in the 24 classes was 418 children. The average number of children per classroom was 17 children (SD= 5.4). Of those, 59 (14%) had parental consent to be tested whereas no response was given by the parents of the remaining children (n=359, 86%). Because the response rate was low, teachers were requested to test all the children with consent, not only the children who appeared to have a risk factor of SNHL. Of the 59 children with consents, 46 children (78%) completed the PAAST whereas 13 (22%) did not because of different reasons such as absence on the testing day (n=4), the child's refusal to be tested (n=3), the child's refusal to complete the test (n=1), the child's sickness at school (n=1) and the teacher having to stop testing the rest of the children in her classroom because she had schoolwork (n=4). The mean number of children tested by each teacher was 2 children (SD= 0.8) with a maximum of 4 children and a minimum of 1 child.

Carers' responses about the risk-factor checklist indicated that out of 59 children, 24 (41%) had at least 1 risk factor of SNHL. Of those, 1 (4%) had 3 risk factors (parental consanguinity, positive family history and parental concern). Another 4 (17%) children had 2 risk factors (3 had parental consanguinity and positive family history whereas the other 1 had parental consanguinity and parental concern). Each of the remaining 19 (79%) children had 1 risk factor (12 had parental consanguinity, 3 had positive family history and 4 had parental concern).

The mean duration required by each child to complete the test was 4.8 minutes (SD=0.7). The mean duration required for the teachers to test all children with consent in their classrooms was 24.3 minutes (SD= 14.6) with a minimum of 4 minutes and a maximum of 50 minutes. The time was calculated by teachers from the first step, where the teacher read the instructions and started to test the first child until she completed testing the last child. The teacher who required 50 minutes to complete the testing tested 4 children and called the children one by one from the classroom to the test room.

7.5 Discussion

One of the main aims of developing the PAAST was to propose its usage for children's hearing screening purposes at homes and schools. The criteria for a screening test included the acceptability of the test by the public and the capability of paramedical or other non-professional individuals to administer the test easily (Hall, 2003). Therefore, the current study aimed to investigate the usability of the PAAST iPad application when administered by parents and teachers. For the purposes of the study, 26 parents used the PAAST to test the hearing of their children (26 children), and 24 teachers used the PAAST to test the hearing of children in their classrooms at schools in Riyadh, KSA. The SUS questionnaire was completed by all subjects and used as the outcome measure to evaluate the usability.

The second aim of this study was to investigate the feasibility of applying a targeted hearing screening at schools in the KSA using the PAAST. For that purpose, 24 teachers from 4 private schools in Riyadh, KSA, distributed the developed checklist of risk factors of SNHL to children's parents to identify children who are at risk of SNHL and used the PAAST to screen the hearing of children in their classrooms (46 children). The feasibility was assessed in terms of the number of teachers who participated in the study, the number of children who performed the hearing screening at schools, the time required to screen the hearing of each child and the overall time required by each teacher to complete the screening. The following sections discuss the findings of the study.

7.5.1 Usability of the PAAST

The usability scores obtained from parents and teachers were close. The PAAST scored 81.3 and 81.5 for parents and teachers respectively. Sauro (2011) reported that a score of ≥68 is considered above average, whereas scores that are <68 are considered below average and thus, improvements should be considered to enhance the performance of the system. According to Bangor *et al.* (2009), who defined an adjective rating for the SUS, any SUS score between 73 and 84 is considered "Good" whereas scores that are ≥85 are considered "Excellent". Based on that, it could be seen that the usability of the PAAST is highly above average and that the score of the PAAST is "Good" (Figure 7.3).

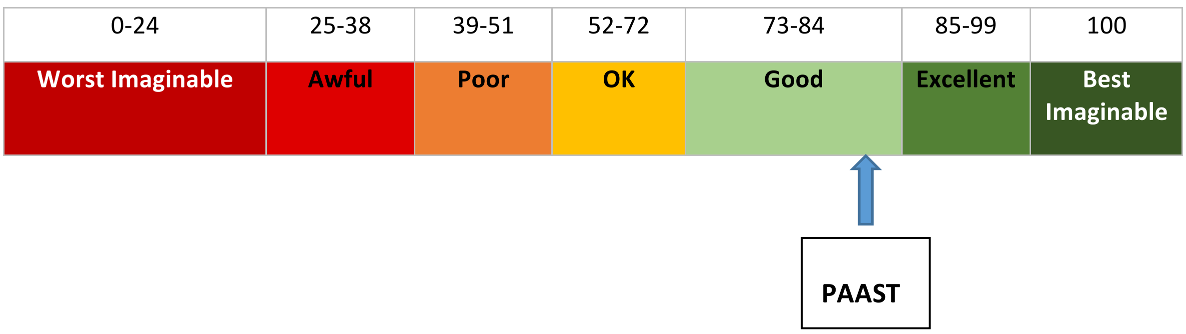


Figure 7.3 The SUS score of the PAAST based on the adjective rating (Bangor *et al.*, 2009).

The SUS itself does not highlight the difficulties or the problems of the system but rather tells how well the users can use the system functions. Therefore, an additional open-ended question was added to allow participants to express their impression about the PAAST. Having such information would help to improve the beta version of the application before its release for public use.

All parents and teachers did not provide written comments about the use of the PAAST. However, some teachers commented verbally about the PAAST and reported the easiness of the PAAST application, the need for such an accessible hearing test and their willingness to use the PAAST with their students in case of suspicion. To get feedback from parents about the PAAST application, two parents were contacted by phone, where they highlighted some positive and negative issues. First, positive feedback was provided about the implementation of the PAAST in an iPad application. This supports the vision about the usefulness of the iPad application in attracting young children to perform the hearing test with no noticeable effort from their parents. It also indicates that the iPad application would enhance the acceptance of people in public to use the PAAST to test children's hearing at home and school, if needed. Second, the parents reported that they liked the use of the visual reinforcement and reported that it helped in gaining the attention of their children as the child seemed to feel proud to know the correct answer. Conversely, parents recommended reducing the frequency of showing the visual reinforcement to reduce the test duration. However, if the reinforcement disappears for a few consecutive responses, the child might feel that s/he is not doing well, her/his performance might be affected negatively, and s/he might refuse to complete the test. Therefore, it was preferred to keep the continuous reinforcements, but reducing its duration from two seconds to one second would be considered. Third, the parents commented positively about the availability of an option to exclude words that are unfamiliar to the child, which supports the idea that the flexibility of the test in terms of excluding the unfamiliar words makes it appropriately usable with most young children.

On the other hand, the parents reported that the large number of choices (14 pictures) might affect the accuracy of the child's response. Although some people might agree with this, the MTT, which has the same number of choices (14 toys) has been used for many years in clinics in the UK and has proved to provide reliable and valid results (Ousey *et al.*, 1989; Harries and Williamson, 2000; Hall *et al.*, 2007). Hence, it seems that the number of choices would not affect the children's performance. Additionally, the parents recommended the use of an "I don't know" sign for the child to select in case s/he was not sure what the presented word was instead of guessing and selecting a random word. The parents' point of view is understood, as they want the child to select only the word that s/he is 100% sure is the correct word. However, the PAAST is developed to estimate the SRT, where the threshold is defined as the point at which the subject can correctly identify the speech stimulus for a specific percentage of the time (70.7% of the time). Hence, at the threshold level, the child is expected to miss the correct words sometimes and to identify it correctly other times. Therefore, it is necessary to ask the child to guess.

Another issue that was raised by the parents is a recommendation to have an option to pause the test. Considering such an option would increase the probability of using the PAAST with young children who are expected to be distracted. However, pausing the test and resuming it at the reached signal-to-noise ratio (SNR) might affect the results by resuming the test at low SNR. However, this issue is worth investigation in a future study to explore the effect of pausing the test and resuming it at a low SNR on the result. Finally, the parents recommended the inclusion of step-by-step instructions within the application. The work on a “Help” section, where the instructions will be available, is already in progress.

No study was found in the literature to report the usability of an e-hearing screening test; therefore, a comparison between the usability of the PAAST and other e-hearing screening tests could not be done. However, the uptake of some other hearing screening tests that are administered through phone such as the digits in noise test was found to be high (Smits *et al.*, 2006; De Sousa *et al.*, 2018). Thus, it seems that e-hearing screening test can be used easily by non-audiologists.

In sum, the PAAST was found to be acceptable and easy to use by parents and teachers. The few points that were raised by parents are minor and will be taken into consideration to improve the performance of the PAAST application.

7.5.2 Feasibility of applying a hearing screening using the PAAST at schools in the KSA

When the usability of the PAAST was investigated, it was found that the PAAST, as the hearing-screening tool, is acceptable and usable by school staff members. However, it was also of interest to investigate whether the process of applying the hearing screening test in school environments is feasible. This is an important investigation to be done prior to proposing an implementation of the preschool hearing screening programme in the KSA. For that reason, teachers from 4 schools in Riyadh, KSA, were recruited to administer the hearing screening to kindergarten children aged 3-6 years at schools.

Out of the total number of teachers who teach kindergarten children at the participating schools, 56% consented to participate. The majority of the participating teachers and headteachers expressed their willingness to have such a test available at schools, at least to use it in case of suspicion. Thus, it seems that the availability of the PAAST would be of assistance for teachers who have concerns about the hearing of any child in their classrooms. The findings indicate that the idea of applying a hearing screening at schools seems to be acceptable by the majority of the teachers.

The aim was to investigate the possibility of applying a targeted preschool hearing screening in schools in the KSA. The idea was to use the risk-factor checklist, which was developed based on the findings of Studies 2 and 3 as the first line of screening, followed by the usage of the PAAST, as the second line, to screen the hearing of children who appeared to be at risk of SNHL. However, because the response rate from the carers of the preschool children was low with only 14% consenting for their children to be tested, teachers were asked to use the PAAST to screen the hearing of all the children with consent. The low response rate does not necessarily indicate that the carers were not interested in the screening test itself because they were not asked about their opinion in applying hearing screening at schools but rather were asked for their permission for their child to participate in a research study. Low response from carers of students to participate in research studies is not uncommon. For instance, the response rate in Fortnum *et al.* (2016), who invited students to participate in a study that aimed to compare the diagnostic accuracy of PTA and HearCheck screener, was 10.4%. In Study 2, when carers were asked about their opinion in applying SEHS, 90% of the carers were happy for their children to be screened for HL at schools. This indicates that applying a hearing screening for children at schools is acceptable to carers.

The findings from the risk-factor checklist revealed that just under half of the children (41%) have at least one risk factor of SNHL. This finding supports the results of Studies 1 and 2 as it indicates that the identified risk factors, particularly consanguinity, are common among children in the KSA (Zakzouk *et al.*, 1993). This highlights the importance of the application of a preschool hearing screening programme in the KSA because it is known that consanguinity is related to late-onset and progressive SNHL.

The mean time required to complete the screening of all children by each teacher seems to be long (24 minutes), especially as the mean number of children tested by each teacher was 2. However, the actual testing time per child was no longer than 5-6 minutes, which is comparable to the finding of Study 5, where the mean duration needed by the young children to complete the PAAST was 5.1 minutes (Chapter 6, Section 6.4.1). The considerably long duration might be explained by the following reasons. First, the teachers were dealing with the PAAST iPad application for the first time; thus, it is expected that they needed time to familiarise themselves with the PAAST. Second, some teachers needed time to call the children from the classrooms to the testing room and back, which made the process longer. This was seen with the teacher who was calling the children one by one from the classroom to the test room; thus, it took her 50 minutes to test 4 children. Although this action increased the duration of the testing, it is understandable that the teachers were avoiding taking the children out of their classrooms unnecessarily for long periods.

Careful regulation needs to be arranged about when to perform the test in order to not interfere with the educational schedule. It is reported in the literature that the SEHS in other countries such as the UK is performed at schools by school health nurses (Bristow *et al.*, 2008). Hence, it seems that it is feasible to perform the screening within school premises. In the KSA, not all schools have school health services, however, the administration of the testing could be done by any staff member at schools. Otherwise, a nurse could be assigned to visit the school and screen the children's hearing similar to the process applied by the school screening program that has been launched recently in the KSA, where nurses visit schools and screen the children for several disorders including HL (Ministry of Health, 2019d).

It is reported in Bristow *et al.* (2008) that it is sometimes challenging to find quiet places at schools to perform the hearing screening. So sometimes the test was administered in cloakrooms, cupboards or even in noisy areas. In the current study, the PAAST was administered in quiet rooms in all the participating schools. However, the number of participating schools was very small, and they were all private schools, which increases the chance of having available quiet rooms (e.g. staff rooms, library etc.). The situation might be different in state schools. Therefore, it is recommended for future study to investigate the feasibility of applying the hearing screening with a larger number of schools including state schools in Riyadh and other regions of the KSA.

In general, bearing in mind the busy environments during the school day, the findings seem promising. The capability of the teachers to test the hearing of the majority of children who have consent (78%) and the cooperation of the majority of the children in performing the test (91%) indicate that applying a hearing screening for kindergarten children in schools in the KSA seems to be feasible.

The study has a strength in terms of the demographical distribution of the participating schools across the city. However, the limitation that all the participating schools were private schools is noted. It is expected for teachers who work in private schools to have more free time than teachers in public schools; thus, they might be more cooperative. Therefore, it is recommended for a future study to investigate the feasibility of applying such a hearing screening in state schools.

7.5.3 Implication of practice

Putting research into practice is not easy. To apply a preschool hearing screening programme at schools, careful protocols including a step-by-step procedure should be investigated and provided. For instance, it appears from the findings of this study that administering the screening for all children by one teacher with no assistance during the school day might not be practical. Therefore, other suggestions could be investigated such as assigning an assistant staff member to collect the children from their classrooms and return them. This would reduce the duration as the teacher will be focusing on testing the children who are supposed to be ready for the test. In addition, the screening could be done during the introductory week of each year, which is the first week of the academic year where no teaching takes place and the children are spending their time playing to get them used to schools' atmosphere. In this case, teachers will not be worried about the children missing their lessons. It is thought that individuals who work at schools such as headteachers, teachers and other staff members are the best placed to provide suggestions about the optimum way and time to apply a preschool hearing screening in schools in the KSA. However, it is important to note here that this study had only investigated the feasibility of applying a hearing screening at schools using the PAAST; it does not propose the implementation of a programme. The implementation of screening programmes should take into account several factors including the economic factor, which are out of the scope of this research.

7.6 Conclusions

The current study aimed to investigate the usability of the PAAST and the feasibility of administering the test at schools in the KSA. The findings of the study are summarised in the followings:

- The PAAST scored high in the SUS when used by parents and teachers suggesting that it is easily usable
- It seems feasible to administer the PAAST at schools in the KSA for hearing screening purposes
- The majority of the school staff members showed an interest in having the application available to be used in case of concern about children's hearing.

Chapter 8 General discussion, conclusions and future work

8.1 General discussion

The ultimate goal of the current PhD project was to provide solutions that would aid the identification of sensorineural hearing loss (SNHL) in children in the Kingdom of Saudi Arabia (KSA) early after onset. The focus was on the sensorineural type of hearing loss (HL) as it is permanent and has serious consequences on the affected children and their families if not identified and addressed as early as possible. Untreated severe SNHL in childhood may result in communication difficulties in adulthood which may increase the risk of mental health problems. Mild degrees of SNHL also have serious adverse effects on children's general wellbeing including speech and language development, social and behavioural wellbeing and academic performance. Therefore, it is essential to identify children with SNHL as early as possible, preferably prior to the appearance of the symptoms or, at least, before it causes disabilities (American Academy of Pediatrics, 2007; American Academy of Audiology, 2011; American speech-Language-Hearing Association, 2016)

Although SNHL is untreatable, rehabilitation and early intervention would reduce the negative impact of SNHL on children's wellbeing and improve the children's quality of life by providing services that help them to cope with the impairment and to benefit effectively from the remaining functioning hearing they have (reviewed thoroughly in Chapter 2, Section 2.2.3). Therefore, childhood hearing-screening programmes are applied in many countries around the world. For instance, a universal newborn hearing screening (UNHS) programme is widely applied and has shown a dramatic reduction in the age of identification (AOI) of children with SNHL (Centers for Disease Control and Prevention, 2003; Wood *et al.*, 2015; Centers for Disease Control and Prevention, 2017). However, UNHS has some limitations as it only targets children with bilateral moderate and worse degrees of SNHL (Public Health England, 2016). Thus, children with mild SNHL and unilateral SNHL would pass the UNHS. Additionally, children who develop late-onset SNHL, which does not manifest at birth but appears later in childhood, and children who acquire SNHL as a result of childhood infectious diseases, ototoxic medications etc. would not be identified early after onset if no further hearing screening were available. Therefore, applying periodic childhood hearing screening following to the UNHS has been recommended (American Academy of Pediatrics, 2016). In high-income countries (HICs) such as the United Kingdom (UK) and the United States of America (USA), school entry hearing screening-programmes are applied for that purpose (Sekhar *et al.*, 2013; Fortnum *et al.*, 2016).

In the KSA, the prevalence of SNHL in children is high, as reported in several studies (Abolfotouh *et al.*, 1995; Al-Abduljawad and Zakzouk, 2003; Al-Rowaily *et al.*, 2012; Alharbi and Ahmed, 2015). This is not surprising, particularly because of the reported high practice of consanguineous marriages in the country (Zakzouk *et al.*, 1993), which is known to be related to late-onset and progressive SNHL (Zakzouk *et al.*, 1995; Schrijver, 2004; Shawky *et al.*, 2013). Despite of this, there was no childhood hearing screening programme in the KSA until 2016, a few months after this PhD project was started, when the Saudi Ministry of Health (MOH) launched a UNHS in the KSA (Ministry of Health, 2016). Afterwards, in 2018, with the huge evolution in the country in many areas including the health services, the Saudi MOH launched a school screening programme that targets year 1 and year 4 students and screens them for several disorders including HL. The programme is currently in its second phase, where only 50% of the students enrolled in year 1 and year 4 are targeted (Ministry of Health, 2019d). This is a commendable improvement in children' health services in the KSA and it is recommended that it be continued.

In line with the evolution of health services in the KSA, particularly the initiation of childhood hearing screening programmes, this PhD project was interested in providing an accessible and usable hearing-screening tool that could be used to screen for SNHL in children. The tool developed can be used to screen the hearing of children immediately in case of suspicion as it does not require special audiology equipment or previous experience of audiology services. It was also of interest to investigate the feasibility of applying a targeted hearing-screening programme that targets kindergarten children who are at risk of SNHL in order to identify SNHL in children early after onset. It seems reasonable to apply a targeted preschool hearing-screening programme in addition to the school screening programme because children in the KSA join year 1 at the age of 6 years. Thus, there is a big time-gap between the UNHS and the school hearing screening. Therefore, it is thought that it would be useful to apply a targeted hearing-screening programme at a kindergarten age in order to identify children with SNHL early after onset. It is believed that applying a targeted hearing-screening programme at this age, rather than universal screening, is more practical, particularly as a universal school screening will follow it during the first year of school.

8.1.1 Age of identification of SNHL in children in the KSA and the characteristics of the affected children

In order to assess the effectiveness of the UNHS programme in reducing the AOI of SNHL in children in the KSA, it is necessary to have information about the AOI before the application of the programme in order to compare it with the outcome. No study has looked previously at the AOI of

SNHL in children in the KSA. In addition, no recent study looked at the risk factors of SNHL in children in the KSA. Providing information about the AOI of SNHL in children would help to show the situation in the KSA and would form a database that could be used to evaluate the effectiveness of the newly applied UNHS. On the other hand, providing information about the risk factors of SNHL in children in the KSA would help in identifying the at-risk children if targeted hearing screening is applied. Therefore, this PhD project looked at the AOI of SNHL in children in the KSA and the characteristics of the affected children.

Two studies (Studies 1 and 2) were conducted for that purpose. Data were driven from four public and private audiology clinics in two main cities of the KSA; Riyadh and Dammam. The findings showed that the AOI of SNHL is high (around 3 years old) although some of the data were driven from hospitals that were independently applying newborn hearing screening to infants born in those hospitals (2 hospitals applied UNHS, 1 applied targeted hearing screening to infants admitted to the newborn intensive care unit, and 1 had no hearing screening). However, it is worth mentioning that even in the hospitals that apply UNHS, it is expected that some children were not screened for HL at birth either because they were born before the hospitals started the hearing screening or because they were born in other hospitals. Hence, the data included children who were screened at birth and children who were not. Besides, the generalisation of the findings to the whole population of children in the KSA might not be appropriate because the data were taken only from two cities in the country. However, it is expected for the AOI in other cities and rural areas of the countries to be higher because Riyadh and Dammam host main hospitals that have audiology clinics, whereas the audiology clinics in other areas are limited.

This finding of the high AOI has two possible explanations. First, the high AOI is caused by the absence of newborn hearing screening. Second, the high AOI is caused by late-onset, progressive or acquired SNHL. Both explanations are partially true, at least to some children. If the first explanation is the main cause, then it is expected that the application of the UNHS would reduce the AOI of SNHL similar to its results in other countries (Dalzell *et al.*, 2000; Wood *et al.*, 2015). This could be investigated in future work by conducting a study to look at the AOI of SNHL in children in the KSA a few years after a full coverage of the UNHS is applied. However, it is expected that the second explanation also plays a role in the observed high AOI of SNHL, particularly as the findings revealed that consanguinity and history of chemotherapy are possible predictors of SNHL in children in the KSA. Consanguinity is related to the presence of late-onset and progressive SNHL, which does not manifest at birth (Zakzouk *et al.*, 1995; Shawky *et al.*, 2013) and exposure to chemotherapy occurs postnatally. Hence, there is a percentage of children who developed SNHL at a later age that contributed to the findings.

The findings also revealed that positive family history of SNHL, history of brain pathologies/tumours and parental concern about children's hearing are potential predictors of SNHL in children in the KSA. Parental concern about children's hearing should be considered seriously because it seems that parents may have a concern about their child's hearing, but because of the low awareness of HL and its consequences, they do not respond to that concern. Raising the awareness of parents about HL and its consequences is highly recommended. Parents should know that any concern about the child's hearing should be taken seriously, and they should seek immediate help. Raising the awareness of general practitioners (GPs) about the importance of taking parents' concern about a child's hearing seriously is of equal importance. This is because it was highlighted in Study 2 that in some cases GPs did not refer the children for a hearing assessment regardless of the expression of the parents' concerns about their children's hearing. The role of GPs in identifying children with HL should not be underestimated, particularly in a country like the KSA where audiology clinics are only located in the hospitals in the main cities and where the health system requires patients to first visit GPs for any health complaints. In the UK, in an area where school entry hearing screening was not applied, most of the children who were identified with HL were referred by their GPs (Bristow *et al.*, 2008), which highlights the important role of GPs in identifying children with HL. The lack of audiology clinics in the KSA is another factor that was highlighted by parents as contributing to the high AOI of SNHL as the accessibility to audiology clinics is not easy and requires a long referral process.

Based on those findings, it seems useful to provide an easily administered hearing-screening test to be used by parents and/or GPs in case of suspicion about a child's hearing. This would partially solve the problem of difficulty in accessing audiology clinics since GPs could screen the child's hearing and refer only children who fail the screening test. Additionally, parents could use it to screen the hearing of their child in case of suspicion and seek help if the child fails the test. Therefore, this PhD project worked on the development of a hearing screening tool that is easily accessible and usable by individuals who have no experience in audiology. This hearing-screening tool could be used to screen children's hearing informally in case of suspicion and/or formally in childhood hearing-screening programmes such as those that target children who are at risk of SNHL.

8.1.2 The development of the Paediatric Arabic Auditory Speech Test (PAAST)

It is known that individuals with SNHL complain mainly about the reduced ability to understand speech in the presence of background noise. It is recommended that this ability is assessed directly using speech-in-noise (SIN) tests rather than by routine pure tone audiometry (PTA). This is because

PTA is an assessment of the ability to detect pure tones, which is a measure of audibility, whereas the ability to understand speech is not only affected by the audibility of the sound but other psychoacoustic factors, such as frequency selectivity, temporal resolution and cochlear compression, also contribute to it (Moore, 2003). SNHL is known to cause an impairment of one or more of these factors, which in turn distort the received speech. So, even if the speech is audible, the distortion effect of SNHL makes it difficult for the individual to understand it (Plomp, 1978). The situation becomes worse with the presence of background noise. Therefore, PTA is not an ideal test to measure the ability to understand SIN, whereas SIN tests, which use speech materials and present it in noise, are.

The current PhD work developed a SIN test for children in the KSA, called the Paediatric Arabic Auditory Speech Test (PAAST). The PAAST is inspired by the McCormick Toy Discrimination test (MTT), which is a reliable speech test that suits children from the age of 2 years onward and is used widely in the UK to assess the ability of children to understand speech (McCormick, 1977; Hall *et al.*, 2007; Lovett *et al.*, 2013). The PAAST measures the speech reception threshold (SRT) of children in quiet and in noise by assessing their ability to discriminate between pairs of words that share the same vowels. The PAAST was developed carefully to meet the criteria of screening tests in order to be used for screening purposes. It was implemented in an iPad application that runs the test automatically. This is believed to ease the access to the test and to allow individuals who have no experience in audiology to use the PAAST easily. Hence, parents, GPs, teachers etc. could use it to screen children's hearing in case of suspicion. The words of the PAAST were selected carefully to suit children from the age of 3 years in order to be used appropriately with young children. The familiarity of the words to children in Gulf countries and some Levant countries was considered due to the lack of Arabic SIN tests. This would widen the use of the PAAST in Arab countries other than the KSA.

The development of the PAAST is also believed to add to the clinical diagnostic and rehabilitative services provided to children with SNHL in the KSA. The inclusion of a SIN test in the audiological test battery of children with SNHL is essential to assess the functionality of their hearing properly and to provide them with the best rehabilitation that fits their individual needs. No such test is used in audiology clinics in the KSA because there is no available SIN test that suits children in the KSA. Until 2018, after the data collection of this PhD project was completed, no SIN test for Arabic-speaking children was published. In 2018, an Arabic version of the words-in-noise test (WIN) was published (Rahman, 2018). The test measures the children's speech reception at fixed intensities and targets Egyptian children. This test is not suitable for Saudi children because the words used are in an Egyptian dialect that would not be familiar to children in the KSA, particularly those of a young age.

At this stage, the background noise that was used for the PAAST is the speech-shaped noise that matches the long-term average speech spectrum of the speech material. This type of noise is standard and widely used in SIN tests (Plomp and Mimpen, 1979; Smits *et al.*, 2004; Ozimek *et al.*, 2009; Lovett *et al.*, 2013; Smits *et al.*, 2013; Vaez *et al.*, 2014; Potgieter *et al.*, 2016; Semeraro *et al.*, 2017; Yuen *et al.*, 2019). It might be argued that speech-shaped noise is not realistic unlike babble noise, which is composed of speech of multi-talkers; however, noises that interfere with conversations and cause difficulty in understanding speech in real-life listening situations are not always speech. It could be noise from equipment, vehicles etc. Besides, babble noise that consists of multi-talker speech becomes less speech-like as the speech become unidentifiable. However, it is recommended for other types of noise such as babble noise to be used in the future with the PAAST and to investigate the difference between the performances of children in different types of noise.

In Study 3, the speech material of the PAAST was equalised for its intelligibility (the homogeneity of the words was enhanced). This step was necessary in order to ensure that at a given signal-to-noise (SNR) level, all words are equally identifiable. Equalising the intelligibility of the words enhances the accuracy of the test in estimating SRTs by reducing the estimation errors. In order to equalise the intelligibility of the words of the PAAST, the psychometric function (PF) of each word was fitted, and the location of each word was assessed and manipulated to match the average locations of the words (see Chapter 4, Study 3 for more details). The literature showed that previous studies equalised the intelligibility of their speech material to match a pre-identified percentage of correct responses (50%, 70.7%, etc.), which matches the adaptive procedure they aimed to use in the test (Plomp and Mimpen, 1979; Summerfield *et al.*, 1994; Smits *et al.*, 2004; Ozimek *et al.*, 2009; Vaez *et al.*, 2014; Potgieter *et al.*, 2016; Semeraro *et al.*, 2017; Yuen *et al.*, 2019). However, in Study 3, the equalisation was made to match the average location of the words (which is the steepest point on the x-axis of the PF curve). This was believed to allow for using the equalised words in any adaptive procedure or even in tests which aim to measure the percentage correct at fixed SNRs without worrying about the procedure used in the process of the intelligibility equalisation.

The speech materials were utilised in a 1-up/2-down adaptive procedure, which corresponds to an SRT of 70.7% correct responses. At this stage, the PAAST was implemented in an iPad application to meet the recommendation of an easily accessible hearing-screening test. The adaptive procedure that is run automatically by the iPad application meets the requirement of an easily usable hearing-screening test. Besides, the interface was designed to be child-friendly with a smiley face popping-up after each response in order to attract the attention of young children for a longer time. The reliability of the PAAST implemented in an iPad application was assessed against its reliability when administered through MATLAB software with normal-hearing (NH) Arabic-speaking adults (Study 4) in order to ensure that the iPad application was functioning as expected before using it with children. This also provided information about the performance of NH Arabic-speaking adults in SIN tests, which was not available before in literature.

The PAAST was then tested on NH Arabic-speaking children of two age groups (3.0 years to 5 years and 11 months and 6.0 to 12.0 years old) (Study 5). An age effect on the children's performance was found similar to that reported in literature for children speaking other languages (Fallon *et al.*, 2000; Johnson, 2000; Litovsky, 2005; Stuart, 2005; Johnstone and Litovsky, 2006; Stuart *et al.*, 2006; Talarico *et al.*, 2007; Stuart, 2008; Vaillancourt *et al.*, 2008; Lovett *et al.*, 2012; Sheikh Rashid *et al.*, 2017; Vickers *et al.*, 2018). As age increases, the performance of children in SIN tests improves. The reliability of the PAAST, when tested on children, was high. The reliability measurements of the PAAST indicate that the PAAST showed better test-retest reliability than the MTT in noise (Summerfield *et al.*, 1994; Lovett *et al.*, 2013). Since the PAAST was found to be reliable in measuring children's SRT in noise, preliminary normative data were provided. Normal ranges were calculated and a cut-off value for normal results was identified for both groups of children (young and older children). Additionally, the test duration was estimated in order to assess the practicality of using the PAAST for screening purposes, which require the use of quick tests and in clinical settings, which is usually busy. Children were able to complete one run of the PAAST in around 5 minutes, which is expected to be reasonable both in clinics and screening settings.

The PAAST was then performed by Arabic-speaking children who have SNHL in order to investigate the ability of the PAAST to differentiate between NH children and children with SNHL (Study 6). The findings showed that the PAAST can differentiate, to some extent, between NH children and children with SNHL. However, some overlap was observed between the two groups at the borderline point, where some children with SNHL showed normal SRTs (based on the determined cut-off value). This indicates that those children might not be affected severely by the distortion component of the SNHL; thus they have less difficulty in understanding SIN. It is also possible that those children use coping strategies that help them function better in SIN situations. On the other hand, the overlap could be drawn back to the determined cut-off value, which was based on findings from a small sample size (Study 5). Thus, the determined cut-off value might not be accurate. Although the findings are promising, it is recommended for future studies to conduct a large-scale normative data study and to investigate different cut-off values in order to determine the optimal cut-off value that would best differentiate between NH children and children with SNHL.

If the PAAST will be used for children's hearing screening purposes, it would possibly pick up children who have any type of HL including children with temporary CHL. This would lead to high referral rate because it is known that otitis media with effusion (OME), which could result in temporary CHL, is common in children of young ages. Prevalence of OME among a sample of 1488 school children in the KSA was reported to be as high as 7.5% (Humaid *et al.*, 2014). It was also reported in the same study that the presence of OME was strongly associated with the presence of CHL and with age less than 8 years. In a study that aimed to investigate the application of a hearing screening (PTA) for preschool children aged 3-7 years (n= 6288), 86% of the children identified with HL were found to have CHL as a result of OME (Wu *et al.*, 2014), indicating that the majority of children had temporary CHL. Picking up children with temporary CHL is not recommended if the screening aims to identify children with SNHL. This is because it would increase the referral rate of cases that are not of interest of this particular screening, which in turn would unnecessarily increase the load of work in audiology clinics. It is important to note here that the author is not underestimating the negative impact of CHL on children's general health, but as mentioned in Chapter 1, SNHL is irreversible and has serious negative impact on children's health and wellbeing if not identified early after onset whereas it is known that most cases of CHL in children are temporary and treatable.

It is believed that the method used in Study 6 (see Chapter 6, Section 6.6.2), where the testing started at a loudness level that is comfortably loud to each child, increased the sensitivity of the test to SNHL and reduced it to CHL, at least in principle. This is because it is known that children with CHL suffer mainly from reduced audibility and that increasing the loudness can compensate for the disadvantage of low SNR to individuals with CHL (Nia and Bance, 2001). So, to reduce the sensitivity of the PAAST to CHL, it is recommended to use the same method used in Study 6 if the PAAST will be used to screen for SNHL. Otherwise, if one is interested in identifying children who have HL of any type, the test could be performed by starting the testing at a fixed loudness level that is comfortable to NH children and in this case it is expected for the test to be sensitive to CHL as well as SNHL. However, the sensitivity of the PAAST to CHL is not known as it was not investigated in this research. It is therefore recommended for a future study to look at the sensitivity of the PAAST to CHL in children.

The findings of Studies 4-6 revealed that the PAAST can reliably measure the SRT in noise of Arabic-speaking adults and children and, to some extent, it can differentiate between NH children and children with SNHL. These findings are promising as they indicate that with some improvements, the PAAST can be released to be used in clinics. This is expected to add to the audiological services provided to children with SNHL in the KSA and other Arab countries, at least in Gulf countries. The benefit of the PAAST is not limited to children who have SNHL; children who are expected to have difficulties understanding SIN due to other disorders can also benefit from the PAAST. For example, children with an auditory processing disorder are known to have difficulties understanding SIN despite their normal audiograms. The availability of the PAAST is valuable in the assessment of those children, particularly as they would pass the hearing screening if PTA was used for that purpose. Besides, the availability of a SIN test within the clinical diagnostic assessment of those children would help in identifying them and quantifying their difficulties, thus aid the rehabilitation services provided to them.

The PAAST can also be used with adults who have severe hearing impairment and/or who use cochlear implants, particularly as no SIN test is available for Arabic-speaking adults. Apparently, the availability of the PAAST is a valuable addition to the audiology services in the KSA. It is worth noting here that the instructions of the PAAST are available in both Arabic and English languages, which allow English-speaking audiologists around the world to use it easily with their Arabic-speaking patients. This adds to the value of the PAAST as it widens its spread around the world.

In addition to the clinical contribution of the PAAST, it also contributes to research as it allows individuals who are interested in studying the performance of Arabic-speaking children to use a test that is not strongly linguistically and cognitively demanding. For example, it could be used in a longitudinal study to look at the effect of age on children's performance without stressing their linguistic and cognitive abilities. Furthermore, the speech stimuli of the PAAST could be presented in other types of noise and in quiet as well. This would open up other areas of research in speech perception abilities of Arabic-speaking children.

8.1.3 The feasibility and usability of using the PAAST as a hearing-screening tool

Based on the findings of Studies 1 and 2, providing a hearing-screening tool that could be easily accessible and usable by individuals who have no experience in audiology, such as parents and teachers, was recommended in order to use it to screen children's hearing in case of suspicion. Additionally, it seems useful to apply targeted preschool hearing screening in the KSA in order to identify children with SNHL early after onset. Therefore, the development of the PAAST considered the implementation of the test in an iPad application, which is expected to ease its accessibility and usage.

Study 7 investigated the usability of the PAAST when administered by parents and teachers. This investigation was essential in identifying possible problems that would appear while the test is used by parents and teachers and to apply improvement if needed. For that purpose, the system usability scale (SUS) questionnaire was distributed to parents of children aged 3-6 years and to teachers of kindergarten children. The questionnaire was filled in after the subject (parent or teacher) administered the test at least to one child. The findings showed that the PAAST has good usability on the SUS (>80) (Bangor *et al.*, 2009). No common problems were identified by the users despite some minor recommendations that were raised by two parents, who were contacted to provide their feedback on the PAAST (See Chapter 7, Study 7 for more details). The recommendations provided will be taken into consideration to improve the PAAST.

Additionally, Study 7 investigated the feasibility of using the PAAST to screen for SNHL in kindergarten children in schools in the KSA. Twenty-four teachers from 4 private schools participated in the study. They distributed the checklist of risk factors of SNHL to parents of children in order to identify children who were at risk of SNHL (the checklist was developed based on the findings of Studies 1 and 2). The responses to the checklist revealed that just below half of the participating children had at least one risk factor of SNHL, which highlights the needs for such a hearing screening programme. The aim was for the teachers to use the PAAST to only screen the hearing of children who have risk factors of SNHL. However, due to the low response to the request for participation in the study from the parents of the children, teachers were asked to test any child who had parental consent to participate in order to be able to achieve the aim of the study (assessing the feasibility of applying a preschool hearing-screening programme in schools in the KSA). The findings revealed that the application of a hearing screening in schools using the PAAST is feasible. This was judged based on the willingness of the headteachers to have the hearing screening done in their schools, the willingness of teachers to administer the test to their students, the acceptance of children to perform the hearing screening in schools and the time needed to complete the testing. However, to apply the hearing screening in schools, it is recommended that a protocol is carefully prepared to ensure that the screening can be done properly without interfering with the learning schedules.

It is worth noting here that the use of SIN tests for children's hearing screening purposes is still new in the field particularly in the context of using e-health. As mentioned in Chapter2, Section 2.4.1, only two SIN tests were found in literature to be developed for that purpose. Those are the Dutch internet-based SIN test and the digit in noise test in Flemish (Sheikh Rashid *et al.*, 2017; Denys *et al.*, 2018). The development of the PAAST, which is a fully automated SIN test for children, would add to the audiology screening services provided to Arabic-speaking children not only in the KSA but in other Arab countries as well. Although the PAAST is not the first SIN test that was developed to screen the hearing of children, the current research has a potential novelty by showing that SIN tests can be used to screen the hearing of young children from the age of 3 years. The two SIN tests that were developed for children's hearing screening were not evaluated on children younger than 5 years old. Thus, to the best of the author's knowledge this research is the first to provide evidence-based information on the possibility and practicality of using SIN tests to screen for HL in young children (as young as 3 years old) either formally for SEHS purposes or informally at home.

8.1.4 Potential applications of the PAAST

There are several potential applications of the PAAST for both screening and diagnostic purposes as shown below.

8.1.4.1 Potential applications for screening purposes

The PAAST can be used to screen children's hearing informally in case of suspicion. Parents, teachers and/or GPs can use it for that purpose. The usability of the PAAST was investigated for that purpose and the findings were promising. However, further investigation needed to be done to determine the optimal cut-off value that differentiate between NH and HL. This should be done prior to releasing the PAAST for public use.

The PAAST can also be used formally in hearing screening programmes such as a targeted preschool hearing-screening programme in schools. In this screening programme the targeted children would be children with SNHL. It is suggested to run the screening at schools because it is the place where children can be easily found in groups. Otherwise, it might be difficult to reach all children at these young ages. The checklist of risk factors of SNHL, which was developed in this research based on the identified risk factors of SNHL in children in the KSA, could be used as the first-line hearing-screening tool to identify the at-risk children. The PAAST could then be used to screen the hearing of the identified children. In this research the PAAST showed an ability to identify children with SNHL, at least children who have remarkable difficulty understanding SIN. The findings showed that it might not be sensitive to mild cases where children do not show high levels of difficulty understanding SIN. To investigate the effectiveness of the PAAST in identifying mild cases of SNHL a future study is recommended to be carried out to estimate the SRT of children with mild SNHL using the PAAST. Additionally, a large-scale study is recommended to be conducted to determine the optimal cut-off value that can be used to differentiate between NH children and children with SNHL. Besides, it is crucial to investigate the sensitivity and specificity of the PAAST prior to its usage for screening purposes. Other issues related to the cost-effectiveness of the screening programme need to be investigated as well.

8.1.4.2 Potential applications in clinic

The PAAST could be used in clinics to aid the diagnosis of some audiological disorders such as SNHL, central auditory processing disorder and auditory neuropathy. Children with central auditory processing disorder, for instance, would have normal PTA thresholds but poor SRT. Thus, the involvement of the PAAST in the auditory test battery would ease the identification of those children as early as possible. Moreover, the PAAST can be used to show the patients and their families the extent of difficulties they face in understanding speech in challenging listening environments, which might increase their awareness of the seriousness of the problem and thus motivate them to commit to the rehabilitation services provided to them (e.g. it would increase the acceptability of using hearing aids).

The PAAST could also be used to assess the effectiveness of auditory prostheses. For instance, audiologists may be interesting in using the PAAST to assess the performance of a child with and without a hearing aid. In this case, it is important to know the repeatability of the PAAST, which is the value that could be used to determine whether a change in the performance of a child is statistically significant or not. Although the repeatability of the PAAST was assessed in this research and showed better results than the MTT, the findings were based on small sample size and therefore it is recommended to reassess this critical value for the PAAST with a large sample of NH children. More importantly, this research did not investigate the test-retest reliability of the PAAST with hearing-impaired children, which might be different than that obtained from NH children. Therefore, it is recommended for a future work to investigate this as it would provide important information for clinical practice.

Additionally, the PAAST could be used to monitor the performance of children overtime. For that purpose, it is recommended to assess the between-session reliability of the PAAST, which would differ than the within-session reliability, which was assessed in this research. To assess the between-session reliability, the performance of the subjects who are tested over different sessions that are separated by a period of time (could range from few hours to few months) is assessed. All these areas of research need to be investigated prior to using the PAAST in clinics for the mentioned purposes. It is also recommended to validate the use of the PAAST with other types of noise such as the babble noise in order to find the best type of noise that can be used with the PAAST to achieve optimal results.

8.1.5 Summary

The current PhD project provided data about the AOI of SNHL in children in the KSA and identified common risk factors of SNHL in children in the KSA. Additionally, an Arabic SIN test (the PAAST) was developed and evaluated. The PAAST, which is implemented in an iPad application, can be used to measure SRTs in noise of Arabic-speaking children as well as adults in the KSA and other Arab countries. It can also be used worldwide by English-speaking audiologists with their Arabic-speaking patients. The PAAST is expected to help in identifying children with SNHL early after onset as it can be used to screen the children's hearing immediately in case of suspicion. Moreover, this PhD project investigated the feasibility of applying a targeted preschool hearing-screening programme in schools in the KSA. The checklist of risk factors of SNHL, which was developed based on the identified risk factors of SNHL in children in the KSA, was used as the first-line hearing-screening tool which was followed by using the PAAST to screen the hearing of the children who were at risk of SNHL. The findings revealed that the application of such a programme is acceptable and feasible.

8.2 Conclusions

The main conclusions of this PhD project are as follows:

- This PhD project provided the first report of the age of identification (AOI) of SNHL in children in the Kingdom of Saudi Arabia (KSA). The AOI of children with SNHL attending audiology clinics in four hospitals in two KSA cities (Riyadh and Dammam) in 2015 was higher than that reported in the USA and UK. This is probably, at least partly, due to the limited availability of newborn hearing screening in KSA prior to 2015, compared to the adoption of universal newborn hearing screening in the USA and UK in the early 2000s. It seems most likely that the AOI was no lower in other regions of the KSA. This study provides baseline information on AOI that could be used to evaluate the recent government policy of promoting universal neonatal hearing screening.
- Potential predictors of SNHL in children in the KSA were identified. These are consanguinity, positive family history of SNHL, parental concern about the child's hearing, history of chemotherapy and history of brain pathologies/tumours. Consanguinity is still practised in the country. Previous research has not identified parental concern about the children's hearing as a potential risk factor of SNHL in children in the KSA in contrast to its identification in other countries such as the USA and UK (Watkin *et al.*, 1990; JCIH, 2007).
- The speech material of the developed speech-in-noise test, called the Paediatric Arabic Auditory Speech Test (PAAST), which was based on the principle of the McCormick Toy Discrimination Test, was recorded and, following adjustments, found to be equally

intelligible. It was also found to be familiar to normal-hearing children from the age of 3 years and hearing-impaired children from the age of 6 years.

- An automated adaptive Arabic SIN test that was implemented in an iPad application, the PAAST, was found to be reliable when tested on normal-hearing adults and children.
- A developmental age effect on the performance of Arabic-speaking children's in SIN tests was found. As the age increased, the children's speech reception threshold (SRT) in noise increased. Preliminary normative data that are age-dependent were provided for two age groups of children (3-6 and >6-12 years old).
- Children with SNHL were generally found to have much worse SRTs in noise than normal-hearing children despite setting the loudness of the speech to be audible.
- The PAAST, which is implemented in an iPad application, showed high usability scores when assessed through the system usability scale questionnaire by both parents and teachers.
- It seems feasible to apply a preschool hearing-screening programme using the developed risk-factor checklist and the PAAST to screen the hearing of kindergarten children in schools in the KSA (at least in private schools).

8.3 Future work

Recommended future work is as follows:

- In the current PhD project, the AOI of SNHL in children in the KSA was investigated in Studies 1 and 2 and found to be high. However, some of the children who were included in the 2 studies were screened for HL at birth and the others were not. The information about the screening for HL at birth was not available to the author; therefore, it was impossible to conclude that the high AOI was caused by the absence of newborn hearing screening or it is high regardless of its presence. Now since UNHS has been launched in the country, it is recommended for a future study to be conducted in the near future to investigate the AOI of SNHL in children who are known to be screened for HL at birth. This would help in evaluating the effectiveness of the UNHS in reducing the AOI of SNHL
- The PAAST was developed to assess SRTs in noise. In the current PhD project, the reliability of the PAAST was assessed with NH children using headphones. No assessment was done with the sound-field presentation. Sound-field presentation is useful in cases where the audiologist is interested in measuring the performance of children who use hearing aids or cochlear implants. Therefore, it is recommended for a future study to consider assessing the reliability of the PAAST in sound-field.

- In the current PhD project, only one type of noise, which is speech-shaped noise, was used. Using this type of noise, the PAAST showed an ability, to some extent, to differentiate between NH children and children with SNHL. However, it is possible that other types of noise would better differentiate between the two groups of children. Therefore, it is recommended for a future study to compare the ability of the PAAST to differentiate between NH children and hearing-impaired children in different noise types.
- It is recommended for a future study to investigate the hearing-impaired about their performance in noisy situations using a self-reported questionnaire and investigate its correlation to their performance with the PAAST.
- It is recommended for a future study to assess the accuracy measurement of the PAAST in terms of the sensitivity, specificity, positive predictive value and negative predictive value. This is to ensure that the PAAST has high sensitivity and specificity before proposing and releasing the test to be used for hearing-screening purposes.
- The current PhD project provided preliminary normative data for children when tested using the PAAST. However, conducting a larger study with a large sample size of children in different year groups is recommended in order to provide age-specific normal ranges and optimal cut-off values.
- It is recommended for a future study to provide normative data for adults. With the lack of Arabic automated SIN tests, the PAAST would be valuable even for adults as it can be used to estimate the SRT of Arabic-speaking adults (e.g. adults with cochlear implants).
- The current PhD project investigated the feasibility of applying preschool hearing screening in schools in the KSA using the PAAST. However, only private schools participated in the study. It is recommended for a future study to duplicate the study in public schools, preferably in rural areas of the country, where the number of students is expected to be larger, and the number of school staff is expected to be smaller.

Appendix A Questionnaire (Study 2)

ERGO reference: 23832

Questionnaire: Characteristics of Children with Hearing Loss in Saudi Arabia

Code: ______________________ Date: ________________

Name: ____________________________________

Please answer all the following questions to the best of your knowledge:

1. **What is your relationship to the patient?**

o Parent

o Sibling

o Other, specify please ____

2. **What is the educational level of the child's mother?**

o No formal qualification

o Primary

o Secondary

o High school

o College

o Postgraduate

3. **What is the educational level of the child's father?**

o No formal qualification

o Primary

o Secondary

o High school

o College

o Postgraduate

4. **Is the mother in paid employment?**

o Yes

o No

5. **What is the approximate monthly income of the family?**

o Less than SR 5000

o 5000 – 10,000

- o 10,000 – 20,000
- o More than 20,000
- o Prefer not to say

6. **Is there any one smoking at home, where the child lives?**

- o Yes
- o No

7. **What is the gender of the child?**

- o Male
- o Female

8. **How old is your child?**

____

9. **What is the child's order in the family?**

____

10. **How many people are living with the child in the same home?**

- o 2 - 5
- o 6 – 10
- o More than 10

11. **Please indicate if your child has had any of the following conditions: "Tick all that apply"**

- o Infections of the mother during pregnancy
- o Low birth weight (less than 1500 g)
- o Asphyxia (deprivation of oxygen after birth)
- o Jaundice (yellowing of a baby's skin and eyes)
- o Admitted in neonatal intensive care unit for more than 5 days
- o Postnatal infections
- o Any syndrome
- o Craniofacial anomalies (malformation of face)

- o Neurodegenerative disorders (disorders to the nerves system)
- o Developmental delay
- o Language delay
- o Head trauma
- o Previously exposed to chemotherapy
- o Other, please specify ____
- o If the child has any syndrome, please specify_____________________________

12. **Has the child previously been given antibiotics or any other medication that you have been told it might cause hearing loss?**

- o Yes
- o No

13. **Are child's parents related?**

- o Yes
- o No

14. **If parents are related, please specify their relation:**

- o Cousins of first degree
- o Cousins of second degree
- o Other

15. **Is there family history of hearing loss?**

- o Yes, please specify the relationship of the affected person to the child

____

- o No

16. **Do you have any concerns about your child's hearing level?**

- o Yes
- o No
- o Not sure

17. **If your answer was yes to the previous question, what raised your concerns? "Tick all that apply"**

- o My child is not responding when called or talked to at normal loudness level.
- o My child is not speaking well.

- o My child tends to raise the volume of television more than normal.
- o My child keeps on asking for repetition of what has been said to him/her.
- o Other people mention that my child may not be hearing well.
- o Other, please specify ____

18. **What made you first visit the audiology clinic? "Tick all that apply"**

- o My concern about my child's hearing
- o Teacher's concern about my child's hearing or performance at school
- o Doctor's referral
- o Other, please specify ____________________________________

19. **If a school hearing screening program is to be implemented in the Saudi Arabia in order to detect hearing loss early in life, would you feel happy for your child's (or his/her siblings' hearing to be tested at school?**

- o Yes
- o No
- o Not sure, I need more information
- o Not applicable (my child has already been diagnosed with hearing loss)

20. **If your answer to the previous question is "no", kindly specify the reason: "Tick all that apply"**

- o It is not beneficial.
- o I prefer to be with my child when he/she undergoes any health assessment.
- o The test results might trigger anxiety.
- o It might make me aware of something that I don't want to know.
- o Other, please specify ____

If your child has already been diagnosed with Hearing loss, please answer the following question:

21. **How old was your child when s/he was first diagnosed with hearing loss?**

____

If your child has been diagnosed with Hearing loss after the age of 3 years, please answer the following questions:

22. **In your opinion what was the reason for diagnosis not being made at an earlier age? "Tick all that apply"**

o There is no audiology clinic in our town.

o We had not been offered referral for hearing test though we have raised our concerns to general practitioner.

o I thought my child was ignoring us when he/her was called until someone advise me to take him/her for a hearing test

o My child's hearing was normal before this age. Symptoms started to appear after the age of 3 years old.

o Other, please specify ____

23. **If you think that your child has had normal hearing before the age of 3 years, please specify what makes you think of that? "Tick all that apply"**

o My child's hearing was tested before the age of 3 years and the results were normal.

o My child had normal language development.

o My child was responding to low sounds but he started not to.

o My child's general behaviour has changed.

o Other, please specify ____

o Not applicable

For Professional Use Only:

Please specify the child's hearing status:

____

____

Appendix B Checklist of risk factors of SNHL in children in the KSA

Study Title: Feasibility and Usability of the Paediatric Arabic Auditory Speech Test (PAAST)

Researcher: Rania Alkahtani **Ethics number:** 45826 **Version:** 1 **Date:** 20-09-2018

Please answer all the following questions to the best of your knowledge:

1. Are the child's parents related?
    - o Yes
    - o No

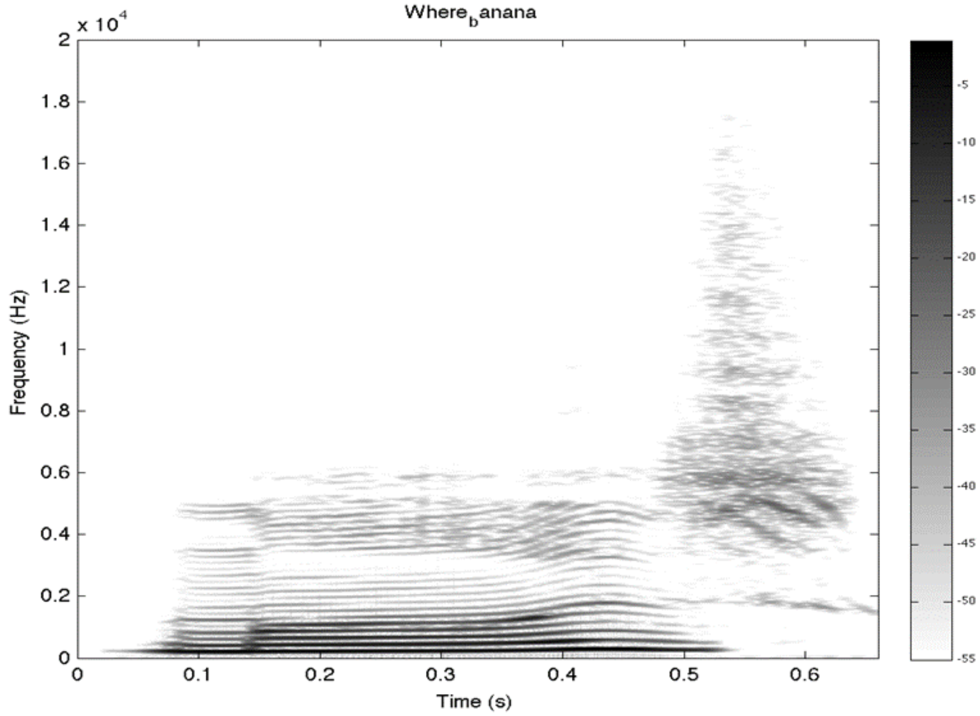
2. Do you have any concerns about your child's hearing level?
    - o Yes
    - o No
    - o Not sure

3. Is there family history of hearing loss?
    - o Yes
    - o No

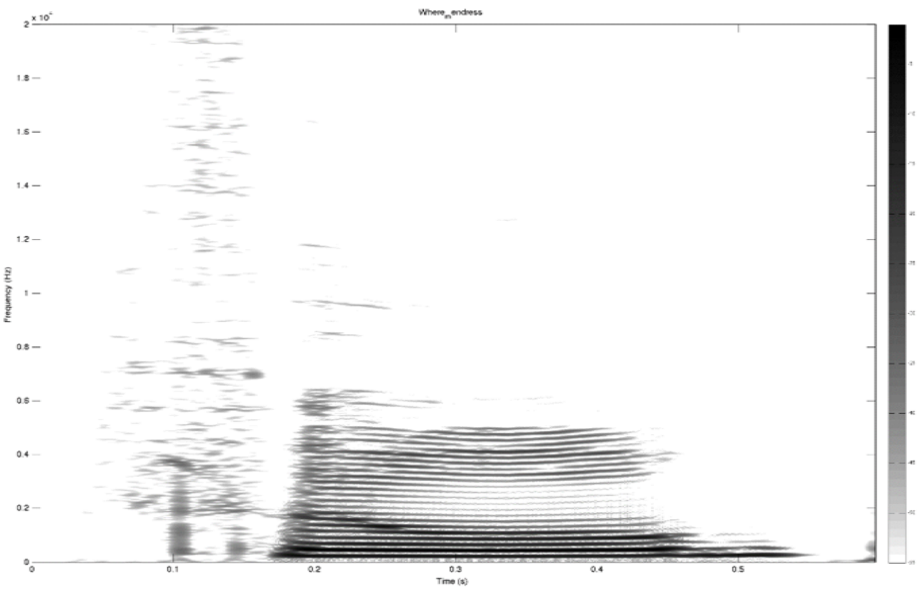
4. Has the child been exposed to chemotherapy?
    - o Yes
    - o No

5. Has the child had brain pathology/Tumour?
    - o Yes
    - o No

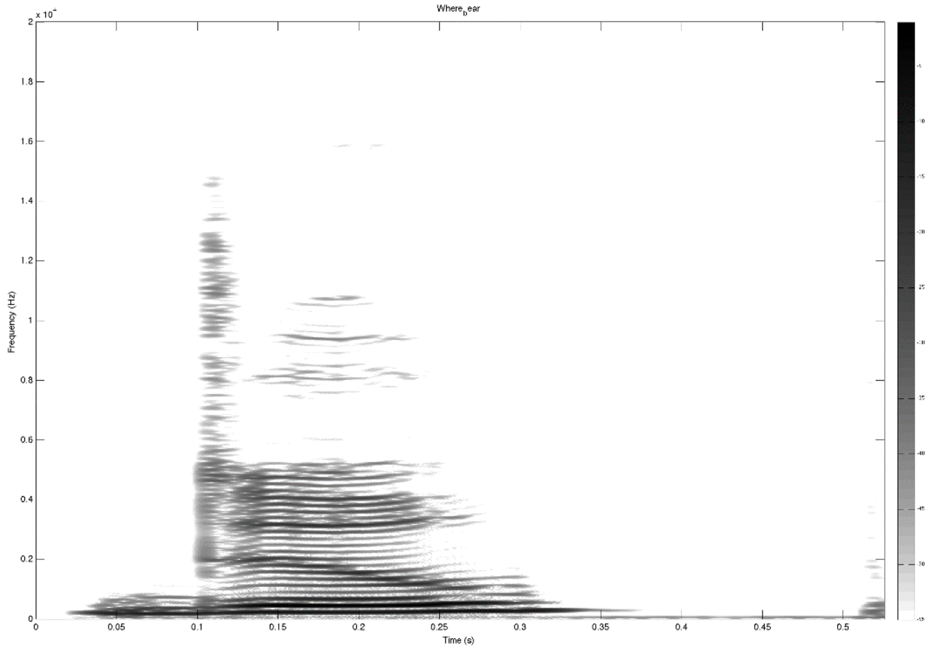
Appendix C Spectrograms of the recorded words



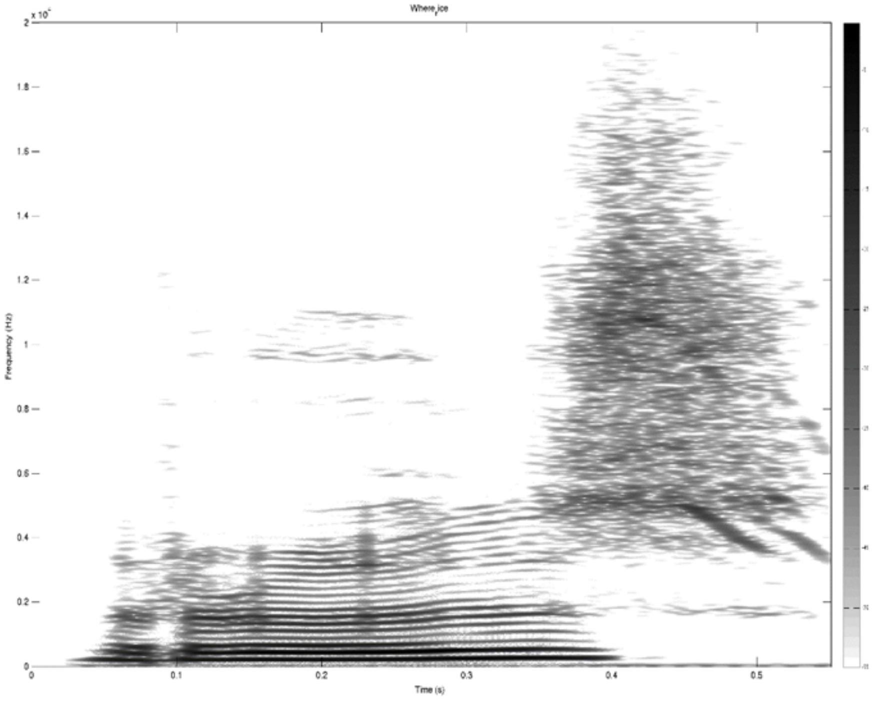
/mɔz/



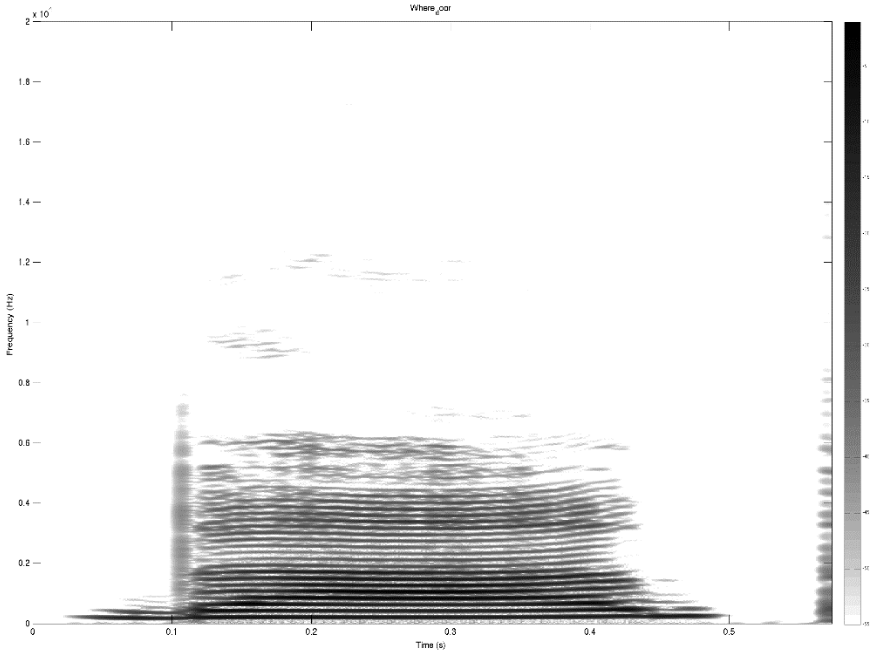
/θɔb/


Frequency (Hz)
Time (s)
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

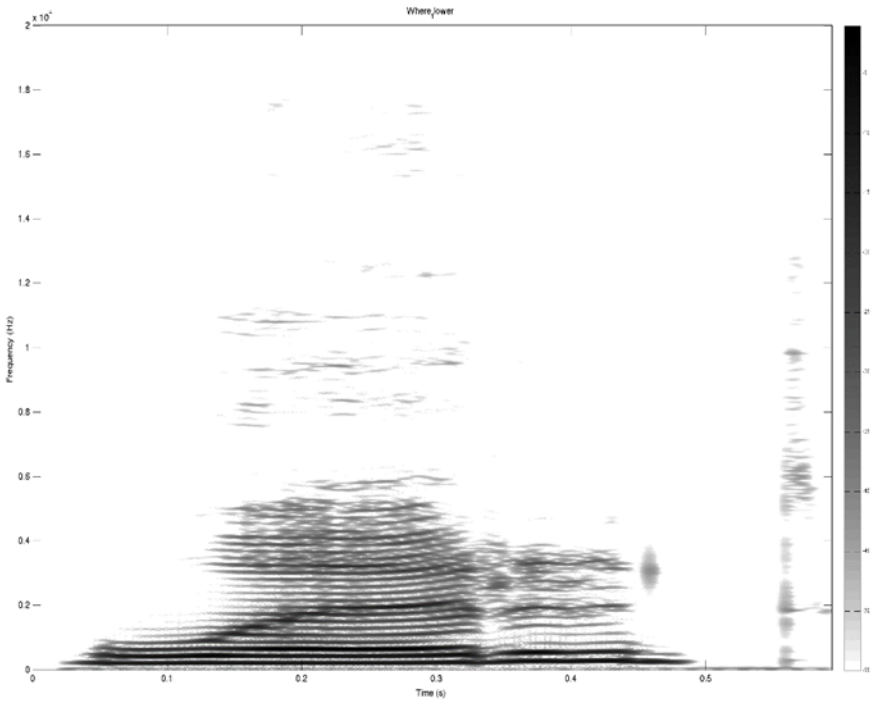
/dob/


Frequency (Hz)
Time (s)
0
0.1
0.2
0.3
0.4
0.5
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

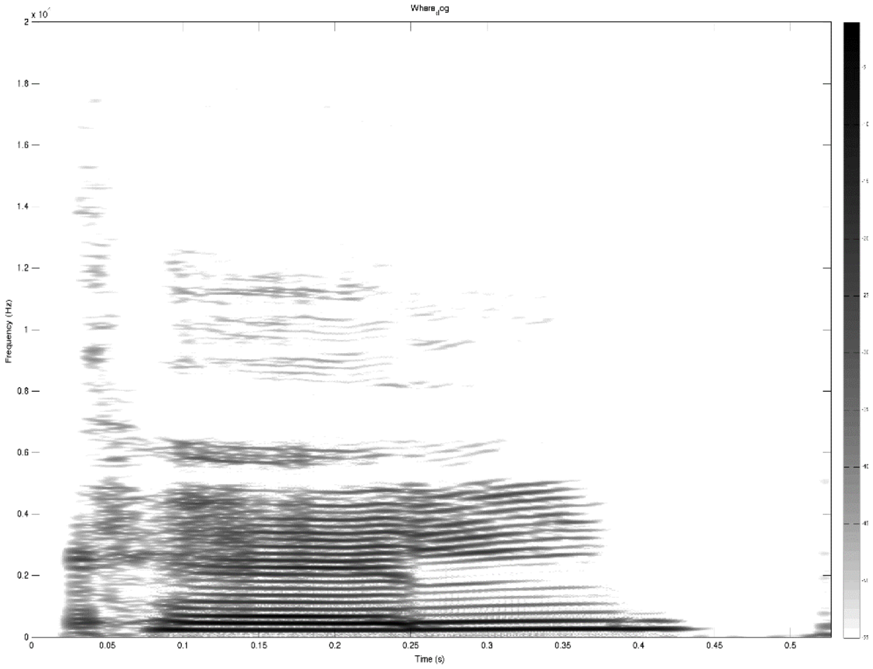
/roz/


Where_door
x 10^4
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Frequency (Hz)
0
0.1
0.2
0.3
0.4
0.5
Time (s)

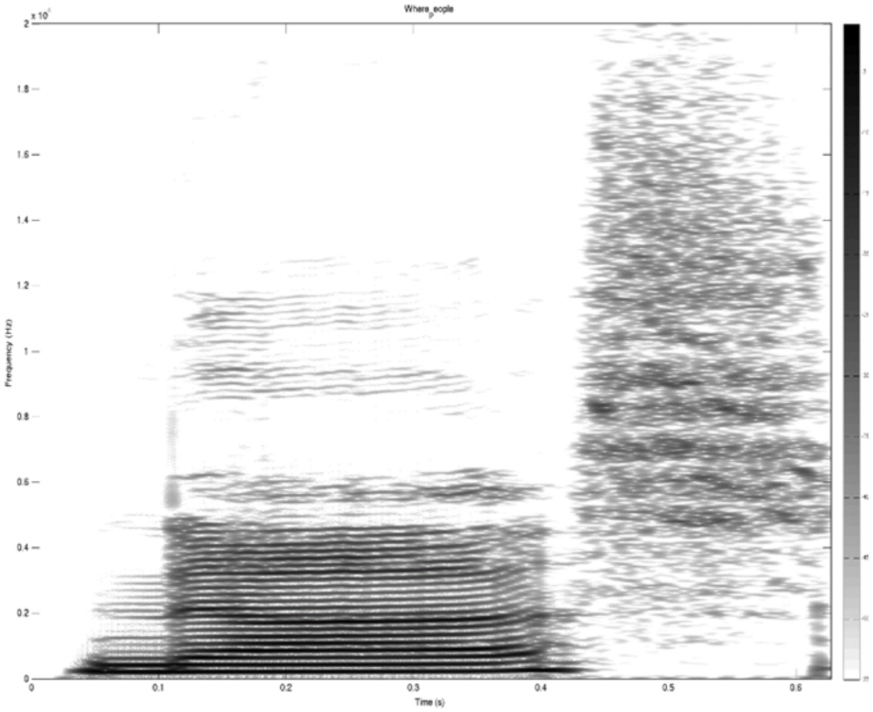
/kʌlb/


Where_flower
x 10^4
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Frequency (Hz)
0
0.1
0.2
0.3
0.4
0.5
Time (s)

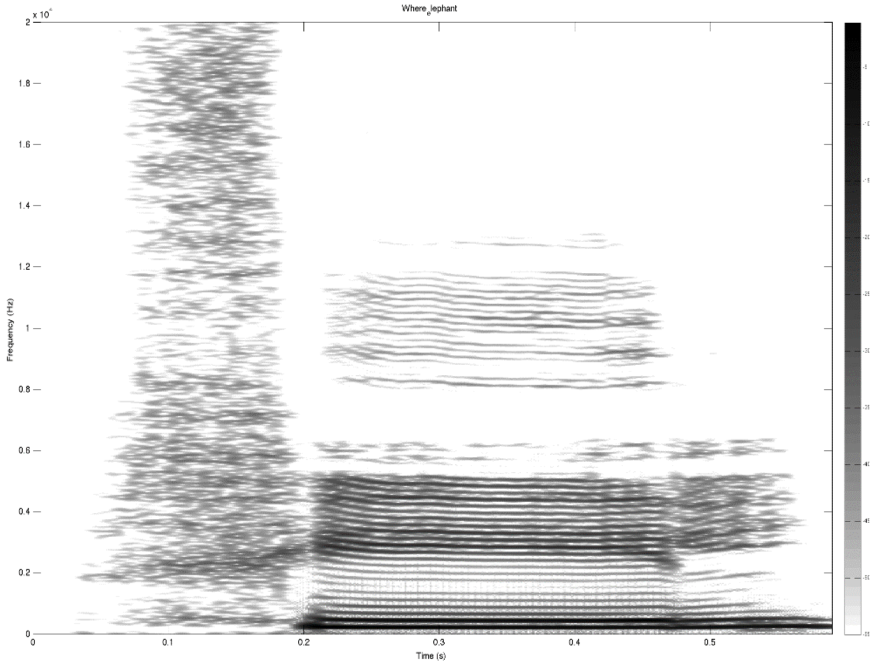
/wʌrd/



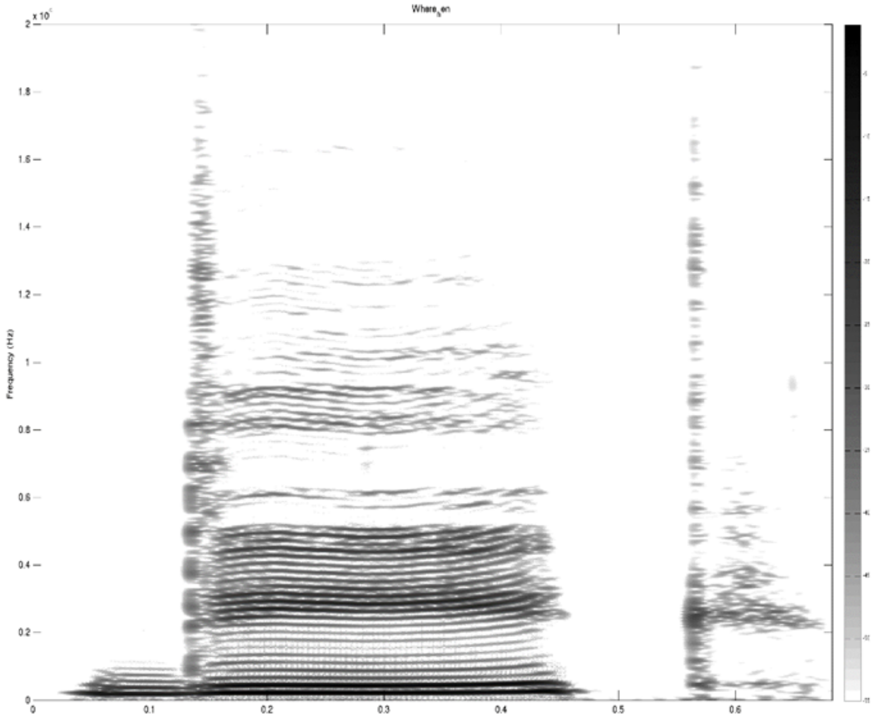
/bab/



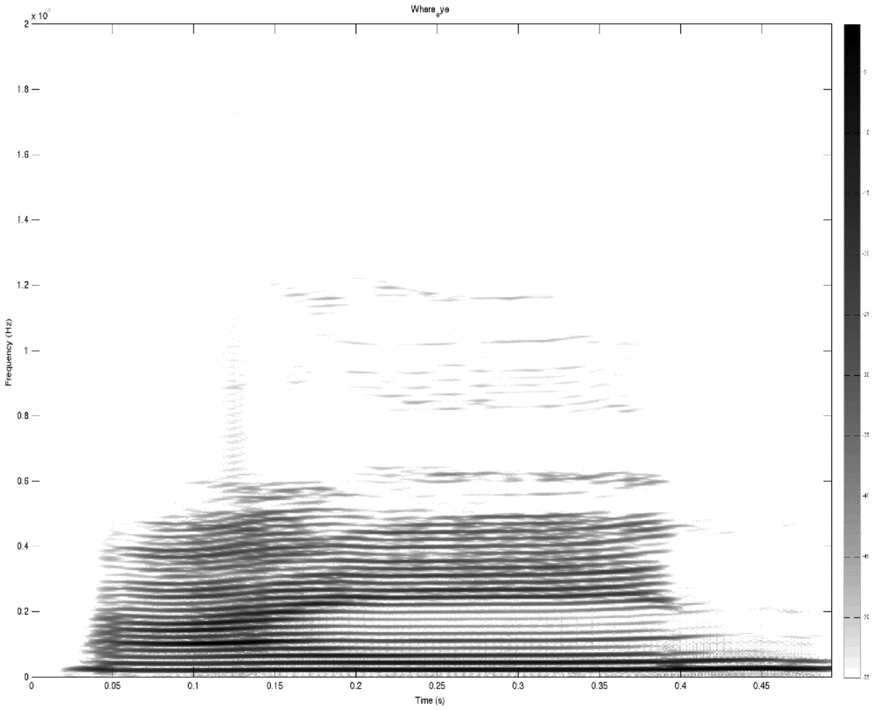
/nas/


Frequency (Hz)
Time (s)
x 10^4
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.1
0.2
0.3
0.4
0.5

/fil/

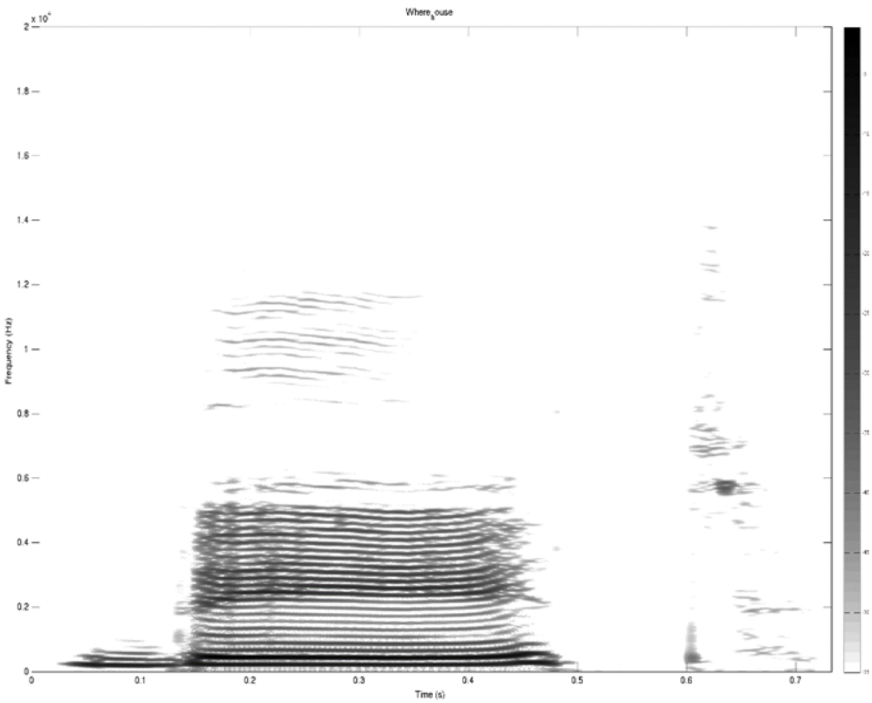

Frequency (Hz)
Time (s)
x 10^4
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0.1
0.2
0.3
0.4
0.5
0.6

/dik/



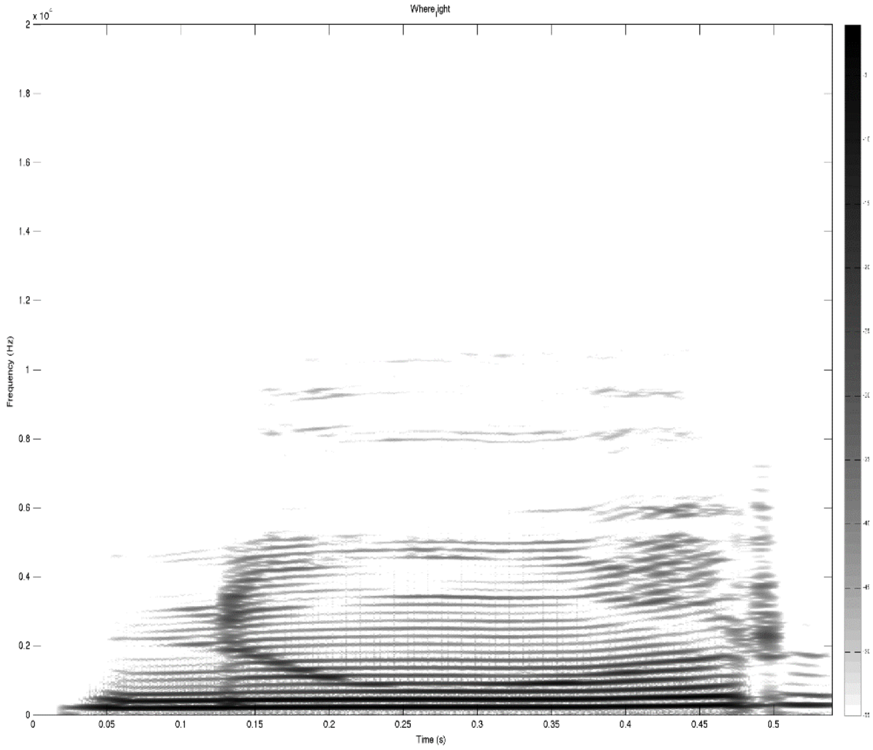
Where$_e$ye
x 10^4
Frequency (Hz)
Time (s)
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

/ʕen/

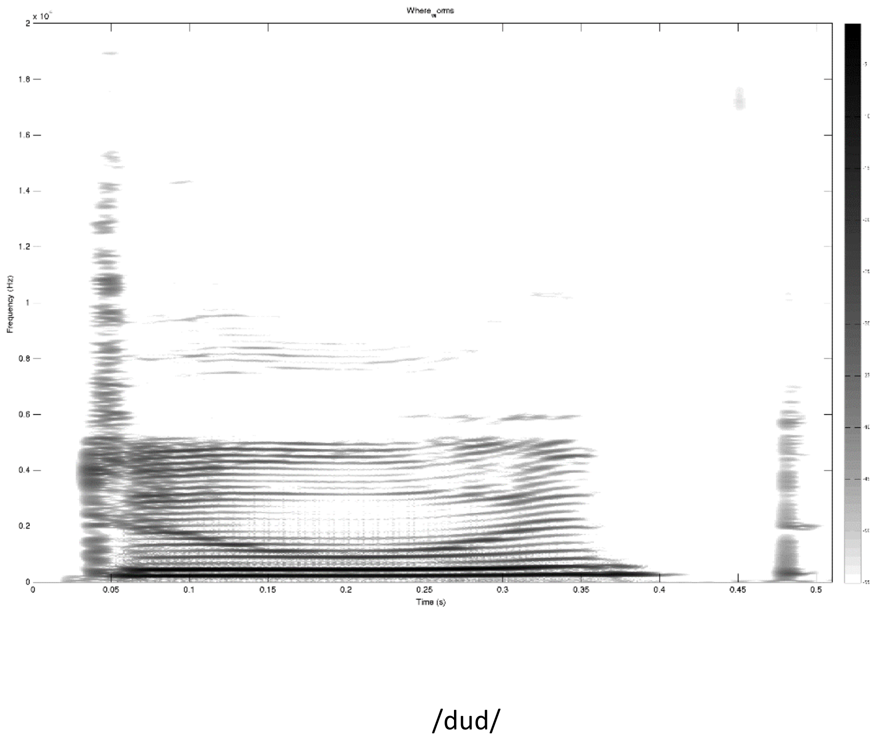


Where$_h$ouse
x 10^4
Frequency (Hz)
Time (s)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

/bet/


Where_right
Frequency (Hz)
Time (s)

/nuɽ/


Where_worms
Frequency (Hz)
Time (s)

/dud/

Appendix D MATLAB code for noise generation

```
% Generation of noise before intelligibility equalisation
clear
close all
% Background info
arabic_toyset = {'banana', 'dog', 'door', 'elephant', 'eye',...
    'bear', 'hen', 'house', 'light', 'mendress', 'rice', 'people',...
    'flower', 'worms'};
arabic_carrier = {'where'};
% Standard rms value
rmsdB_standard = -15; % Sets sounds to be max within any clipping
% Level adjustments
intell_level_adjustments = [-0.5, -1.5, -1.0, 0.5, 2.0, 1.0, 0.0, 2.5, -2.5, -2.0, 1.0, 3.5, 0.5, -3.5]*-1;
% TARGETS
% Read in all sounds, equalise INTELLIGIBILITY levels,
num_toys = length(arabic_toyset);
num_carr = length(arabic_carrier);
stimuli_1 = [];
stimuli_2 = [];
for count = 1:num_toys
    filename_1a = [char(arabic_carrier(1)), '_', char(arabic_toyset(count)), '.wav'];
    [toy_1, fs_1] = audioread(filename_1a);
    rmsdB_toy_1 = 20*log10(rms(toy_1));
    rmsadjust_toy_1(count) = rmsdB_standard - rmsdB_toy_1 + intell_level_adjustments(count);
    toy_1 = 10^(rmsadjust_toy_1(count)/20)*toy_1;
    filename_1b = [char(arabic_carrier(1)), '_', char(arabic_toyset(count)), '.wav'];
    audiowrite(filename_1b, toy_1, fs_1, 'BitsPerSample', 16);
```

```
    rmsdB_toy_1 = 20*log10(rms(toy_1));
    rmsadjust_toy_1(count) = round((rmsdB_standard - rmsdB_toy_1)*10)/10;
    peakdB_toy_1(count) = 20*log10(max(max(abs(toy_1))))
    stimuli_1 = [stimuli_1; toy_1];
end
% CARRIERS
% equalise carrier levels
[carrier_1, fs_c1] = audioread(['carrier_', char(arabic_carrier(1)),'.wav']);
rmsdB_carrier_1 = 20*log10(rms(carrier_1));
rmsadjust_carrier_1 = round((rmsdB_standard - rmsdB_carrier_1)*10)/10;
carrier_1 = 10^(rmsadjust_carrier_1/20)*carrier_1;
peakdB_carrier_1 = 20*log10(max(max(abs(carrier_1))))
% write adjusted carriers
audiowrite('carrier_where.wav', carrier_1, fs_1, 'BitsPerSample', 16);
% MASKERS
stimuli_1 = stimuli_1/sqrt((sum(stimuli_1.^2))/length(stimuli_1)); % normalise using RMS
amplitude
[a,b]= lpc(stimuli_1,6); %generate filter coefficients using LTA spectrum
% generate white noise
noise_duration = 10;
noise_1 = randn(noise_duration*fs_1,1);
% limit the distribution of the noise
peakSDs = 4;
std_1 = std(noise_1)*peakSDs;
peaky_1 = find(noise_1 >= std_1);
noise_1(peaky_1) = std_1;
dippy_1 = find(noise_1 <= -std_1);
noise_1(dippy_1) = -std_1;
% generate speech-shaped noise
speechshapednoise_1 = filter(b,a,noise_1);
```

```
% equalise level of noise (in unlikely event it is required)
rmsdB_noise_1 = 20*log10(rms(speechshapednoise_1));
rmsadjust_noise_1 = round((rmsdB_standard - rmsdB_noise_1)*10)/10;
speechshapednoise_1 = 10^(rmsadjust_noise_1/20)*speechshapednoise_1;
peakdB_noise_1 = 20*log10(max(max(abs(speechshapednoise_1))))
% Plot PSDs of noises
figure
hold on
psd_1 = plot(20*log10(pwelch(speechshapednoise_1)), 'k');
hold off
set(gca, 'XScale', 'log')
% Write noise wav files
audiowrite('arabic_attnoise_where.wav', speechshapednoise_1, fs_1, 'BitsPerSample', 16);
```

Appendix E Psychometric functions

PFs describe the relationship between an observed behavioural performance on a psychophysical task and a given stimulus. It represents how the behaviour change in response to changes in the stimuli. This relationship can be modelled/illustrated using a mathematical function, where a mathematical equation between x and y can be used to show the changes in the variables in response to each other (Kingdom and Prins, 2010).

For speech intelligibility in noise, PFs describe the relationship between the SNR and the proportion of correct responses. As illustrated in Figure D.1, the PF curve represents the SNR on the x-axis and the proportion of correct responses on the y-axis. To fit a PF, observations are made to collect data that ranges from very low SNRs, where it is expected to find no or very few correct responses, to high SNRs, where it is expected to find correct responses to almost all presentations. Once a researcher collects data that represent the performance of the subjects at different SNRs, the data could be used to fit a PF curve using different mathematical functions. If enough data are gathered, the PF curve would be presented graphically in a sigmoid shape (S-shape) (Kingdom and Prins, 2010). Using the PF curve, one can estimate a threshold of interest, which is the lowest level at which the subject can respond correctly for a specific percentage of time.

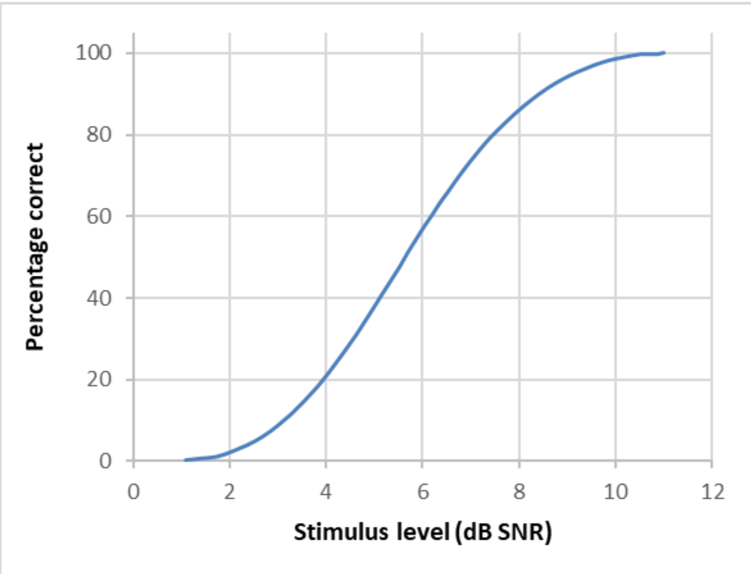


Figure D.1 Graphical representation of a psychometric function

The main features of the PF curve are described in the following (Gescheider, 1997; Kingdom and Prins, 2010):

The guess rate, which represents the lower asymptote of the PF curve. It corresponds to the chance-level performance. At this level, where the SNR (stimulus level) is very low, the subject might find it difficult to identify the speech stimuli, and in this case s/he will need to guess. Therefore, it is expected for the responses at this level to be very low.

The lapse rate, which represents the upper asymptote of the PF curve. It tells how far the scores are from 100%. For NH subjects, it is expected to identify the speech stimuli 100% correctly at a relatively high stimulus level. However, in real practice, a subject might response incorrectly regardless of the high stimulus level, perhaps because of loss of attention or any other reason. Therefore, it is not expected for the responses to reach the maximum at this level.

The slope, which demonstrates the maximum rate at which the response changes as a result of a change in the stimulus level (MacPherson and Akeroyd, 2014). Steep slopes indicate that a small change in the SNR elicits a change in the response. In other words, the steeper the slope, the greater the effect on intelligibility results from a small change in the SNR. On the other hand, for shallow slopes, a small change in the SNR would have no or small effect on the intelligibility, and thus similar responses could be elicited from a range of SNRs. As it appears, steep slope reduces the threshold estimation error.

The location, which represents where the curve is located on the x-axis. The location of a PF depends on many factors such as the type of the employed speech stimuli (Carhart, 1951). For instance, spondee words, disyllabic words and monosyllabic words would result in different presentations of the PF curve. It can be seen in Figure D.2, that the location of the spondee words deviates to the left in comparison to the location of the monosyllabic words. This indicates that spondees words require less stimulus level to elicit the entire PF curve in compare to monosyllabic words. This is also the case with different monosyllabic words where the linguistic characteristics of the words differ and thus result in different locations for each word. It is important to note that the location does not depend on the slope (i.e. different stimuli could have similar slopes but different locations. This is because the slope depends on the stimulus level, whereas the location depends on other factors such as the linguistic feature of the speech stimuli.

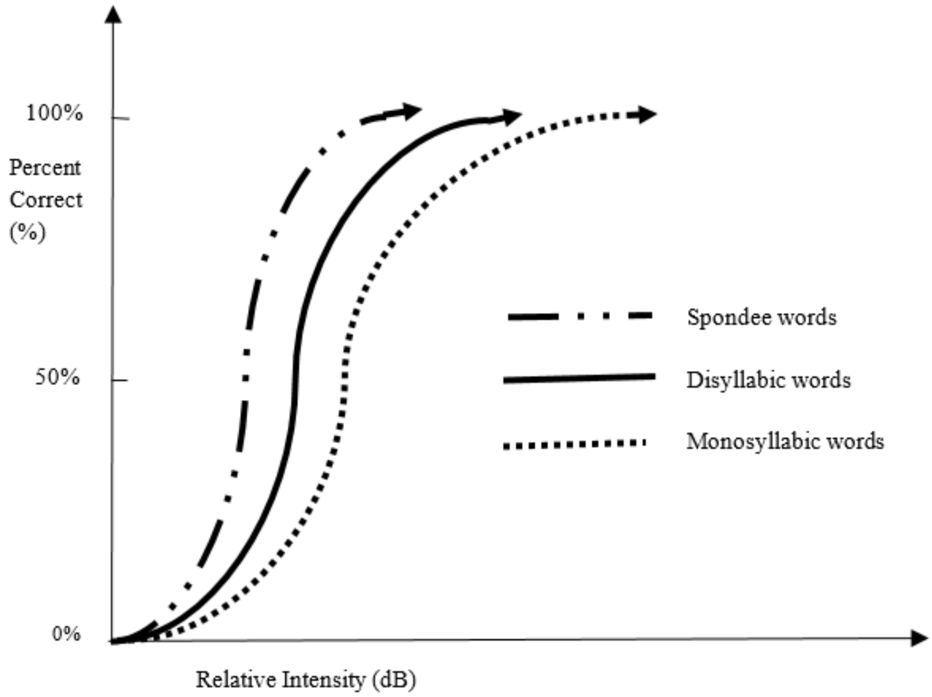


Figure D.2 The psychometric function of different speech materials (taken from Carhart, 1951).

The threshold, which is the SNR at which the response reaches a given criterion (Kingdom and Prins, 2010). The threshold is influenced by the location and the slope, however, when the threshold is close to the location it is less influenced by the slope because the location represents the steepest point on the x-axis of the PF curve. The criterion for determining the threshold is usually set up by the researcher. Some studies defined the threshold as the SNR corresponding to 50% correct responses, and in this case, it is called the SRT_{50} (Smits *et al.*, 2004; Ozimek *et al.*, 2009; Semeraro *et al.*, 2017). Others have different definitions for threshold estimation. For instance, Summerfield *et al.* (1994) defined the threshold as the SNR corresponding to 70.7% correct response. The threshold is read from the PF curve as the SNR for the 50% point, the 70.7% point based on the used definition of the threshold.

Several mathematical functions are available to fit the PFs such as the Logistic, Weibull, Gumbel, Cumulative Normal and Hyperbolic Secant functions. To equalise speech intelligibility, the Logistic function is commonly used to fit the PFs of the stimuli (Summerfield *et al.*, 1994; Smits *et al.*, 2004; Potgieter *et al.*, 2016; Semeraro *et al.*, 2017; Yuen *et al.*, 2019). It is worth noting that for the mathematical calculation of the PFs, the guess rate is not fixed to zero, but it rather fixed at the statistical guess rate of $1/m$, where m is the number of the speech stimuli. This is because although it is expected for the stimulus to be unintelligible at this low level, there is a chance that the subject selects the correct response by guessing and therefore, the guess rate is not expected to be zero. In addition, it is recommended not to fix the lapse rate to 1 but rather to a small number such as

0.01 because, as mentioned previously, subjects might not response correctly 100% of the time when the speech stimulus is expected to be intelligible due to some factors such as loss of attention, sneeze, etc. Therefore, researchers advocate fixing the lapse rate to a small number to avoid bias (Kingdom and Prins, 2010; Klien, 2001; Wichmann and Hill, 2001).

After fitting the PFs using the appropriate function, it is recommended to check how good does the PFs fit the data. This could be done by assessing the probability of deviation (pDev); the goodness-of-fit. This measurement allows for checking the suitability of using the selected function to fit the PFs for a given data. The closer the pDev value to 1, the better the fit. PDev values that are less than 0.05 is an indication that the fit is unacceptably poor (Kingdom and Prins, 2010).

References

Carhart R (1951) Basic principles of speech audiometry. Acta oto-laryngologica 40(1-2): 62-71

Gescheider GA (1997) Psychophysics: The Fundamentals (Third Edition). New Jersey: Lawrence Erlbaum Associates, Inc.

Kingdom F and Prins N (2010) Psychophysics: A Practical Introduction. London: Elsevier Limited

Klien S (2001) Measuring, estimating, and understanding the psychometric function: A commentary. Perception & psychophysics 63: 1421-1455

MacPherson A and Akeroyd MA (2014) Variations in the slope of the psychometric functions for speech intelligibility: A systematic survey. Trends in Hearing 18: 2331216514537722

Ozimek E, Kutzner D, Sęk A and Wicher A (2009) Development and evaluation of Polish digit triplet test for auditory screening. Speech Communication 51(4): 307-316

Potgieter J-M, Swanepoel DW, Myburgh HC, Hopper TC and Smits C (2016) Development and validation of a smartphone-based digits-in-noise hearing test in South African English. International journal of audiology 55(7): 405-411

Semeraro HD, Rowan D, van Besouw RM and Allsopp AA (2017) Development and evaluation of the British English coordinate response measure speech-in-noise test as an occupational hearing assessment tool. International journal of audiology: 1-10

Smits C, Kapteyn TS and Houtgast T (2004) Development and validation of an automatic speech-in-noise screening test by telephone. International journal of audiology 43(1): 15-28

Summerfield Q, Palmer A, Foster J, Marshall D and Twomey T (1994) Clinical evaluation and test-retest reliability of the IHR-McCormick Automated Toy Discrimination Test. British journal of audiology 28(3): 165-179

Wichmann FA and Hill NJ (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. Attention, Perception, & Psychophysics 63(8): 1293-1313

Yuen KCP, Qiu XY, Mou HY and Xi X (2019) The MAndarin spoken word-Picture IDentification test in noise-Adaptive (MAPID-A) measures subtle speech-recognition-in-noise changes and spatial release from masking in very young children. PLoS One 14(1): e0209768

Appendix F Supplementary materials (Study 3)

Table G-1 Locations and slopes and pDevs of all words in Data-1, Data-2 and Combined-data (Session-1)

	1*	**2***	**3***	**4***	**5***	**6***	**7***	**8***	**9***	**10***	**11***	**12***	**13***	**14***
Data 1														
Location	-10.5	-9.7	-10.1	-11.4	-13.2	-12.1	-10.9	-13.2	-8.6	-8.8	-12.0	-14.7	-11.3	-7.4
Slope (β)	0.8	0.5	0.4	0.8	1.0	0.8	0.6	0.7	0.4	0.5	0.4	1.3	0.7	0.5
pDev	0.94	0.80	0.64	0.85	0.80	1.00	0.99	0.49	0.92	0.80	0.69	1.00	0.91	0.90
Data 2														
Location	-10.5	-9.2	-10.6	-11.9	-13.0	-12.3	-10.8	-13.6	-8.6	-10.0	-12.6	-14.5	-12.0	-8.4
Slope (β)	0.9	0.6	0.4	0.8	1.3	0.7	0.5	1.3	0.7	0.5	0.3	1.0	0.9	0.7
pDev	0.90	0.94	0.82	0.89	0.22	0.95	0.86	0.68	0.79	0.37	0.96	0.59	0.78	0.80
Difference between Data 1 and Data 2														
Location	0.0	-1.0	0.4	0.4	-0.2	0.3	-0.2	0.2	0.1	1.2	0.0	-0.1	0.6	1.1
Slope (β)	-0.1	0.1	-0.1	0.1	-0.7	0.1	0.1	-0.3	-0.3	-0.4	0.2	0.0	-0.1	-0.4

Combined Data														
Location	-10.5	-9.5	-10.2	-11.5	-13.2	-12.2	-10.9	-13.3	-8.5	-9.2	-12.1	-14.6	-11.5	-7.6
Slope (β)	0.8	0.5	0.4	0.8	1.1	0.8	0.5	0.9	0.5	0.5	0.4	1.2	0.7	0.5
pDev	0.90	0.89	0.77	0.87	0.75	0.99	0.98	0.53	0.96	0.68	0.86	0.66	0.86	0.94
Adjustment by location	-0.5	-1.6	-0.9	+0.5	+2.1	+1.1	-0.2	+2.3	-2.5	-1.9	+1.1	+3.5	+0.5	-3.4
Adjustment rounded to 0.5 dB	-0.5	-1.5	-1.0	+0.5	+2.0	+1.0	0.0	+2.5	-2.5	-2.0	+1.0	+3.5	+0.5	-3.5

*1:banana, 2: dog, 3: door, 4: elephant, 5: eye, 6: bear, 7: hen, 8: house, 9: light, 10: men-dress, 11: rice, 12: people, 13: flower, 14: worms

Table G-2 Locations and slopes of the words with the lapse rate fixed to 0% and 1% (Session-1)

Word	**1***	**2***	**3***	**4***	**5***	**6***	**7***	**8***	**9***	**10***	**11***	**12***	**13***	**14***
Location (dB SNR) lapse rate 0%	-10.5	-9.5	-10.2	-11.5	-13.2	-12.2	-10.9	-13.3	-8.5	-9.2	-12.1	-14.6	-11.5	-7.6
Location (dB SNR) lapse rate 1%	-10.6	-9.6	-10.3	-11.6	-13.2	-12.2	-11.0	-13.4	-8.6	-9.3	-12.3	-14.7	-11.7	-7.7
Slope (β) lapse rate 0%	0.8	0.5	0.4	0.8	1.1	0.8	0.5	0.9	0.5	0.5	0.4	1.2	0.7	0.5
Slope (β) lapse rate 1%	0.9	0.6	0.4	0.9	1.2	0.8	0.6	1.1	0.5	0.5	0.4	1.4	0.9	0.5

*1:banana, 2: dog, 3: door, 4: elephant, 5: eye, 6: bear, 7: hen, 8: house, 9: light, 10: men-dress, 11: rice, 12: people, 13: flower, 14: worms

Table G-3 Locations, slopes and pDevs of all words (Combined-data, Session-2)

	1*	**2***	**3***	**4***	**5***	**6***	**7***	**8***	**9***	**10***	**11***	**12***	**13***	**14***
Location	-11.4	-11.3	-12.0	-11.2	-11.4	-12.0	-11.1	-10.8	-11.2	-11.1	-11.3	-12.2	-11.8	-10.5
Slope (β)	0.7	0.5	0.5	0.7	0.9	0.7	0.5	0.8	0.5	0.5	0.4	0.8	0.7	0.4
pDev	0.74	0.57	0.73	0.76	0.82	0.83	0.78	0.69	0.78	0.65	0.70	0.77	0.83	0.73

*1:banana, 2: dog, 3: door, 4: elephant, 5: eye, 6: bear, 7: hen, 8: house, 9: light, 10: men-dress, 11: rice, 12: people, 13: flower, 14: worms

Appendix G Instruction sheet

Instruction Sheet (Version 2)

Study Title: Feasibility and Usability of the Paediatric Arabic Auditory Speech Test (PAAST)

Researcher: Rania Alkahtani **Ethics number:** 45826

Overview:

Some children suffer to understand speech especially at noisy environments such as schools, restaurants, funfairs, etc. Those children might be suffering without the notice of people around them. Identifying those children as early as possible allows for intervention to take place and help them to overcome their difficulties as much as possible.

The Paediatric Arabic Auditory Speech Test (PAAST) is a hearing test that is developed to test children's ability to understand speech in the presence of background noise. The test is expected to identify children who are suffering from understanding speech in noisy situations. The test is designed and implemented in an iPad application in order to make it attractive to young children and to make it easily accessible and applicable by both specialized non-specialized persons such as teachers, parents, nurses, etc.

How does the test work?

There are 7 pairs of words (a total of 14 words) represented with pictures. The child is going to listen (through the iPad earphones) to sentences in the form of [where is the "word"] and s/he is required to touch the picture of the word s/he heard on the iPad screen. Once the test is started, it will be run automatically based on the responses of the child. At the end, the result of the test will be displayed on the screen. It is expected for the test to last around 5-6 minutes per child.

General instructions:

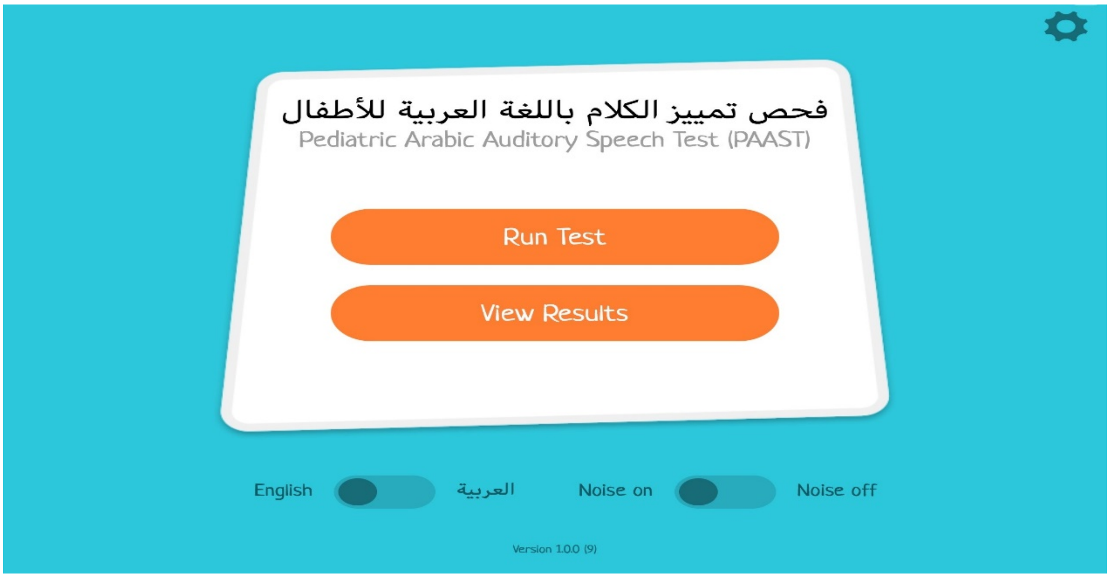
- Please be seated in a quiet room to perform the test.
- It is very important to **<u>use the iPad/iPhone earphones</u>** for this test and not any other kind of headphones/earphones. This is because we tested the output level of the iPad/iPhone earphones and took all the precautions to ensure that the output level of sound is safe for the child's hearing. Using other kinds of headphones/earphones would have some risks so please ensure that you use the recommended earphones.
- **<u>Test each child only once. Do not repeat the test on the same child.</u>**

Appendix G

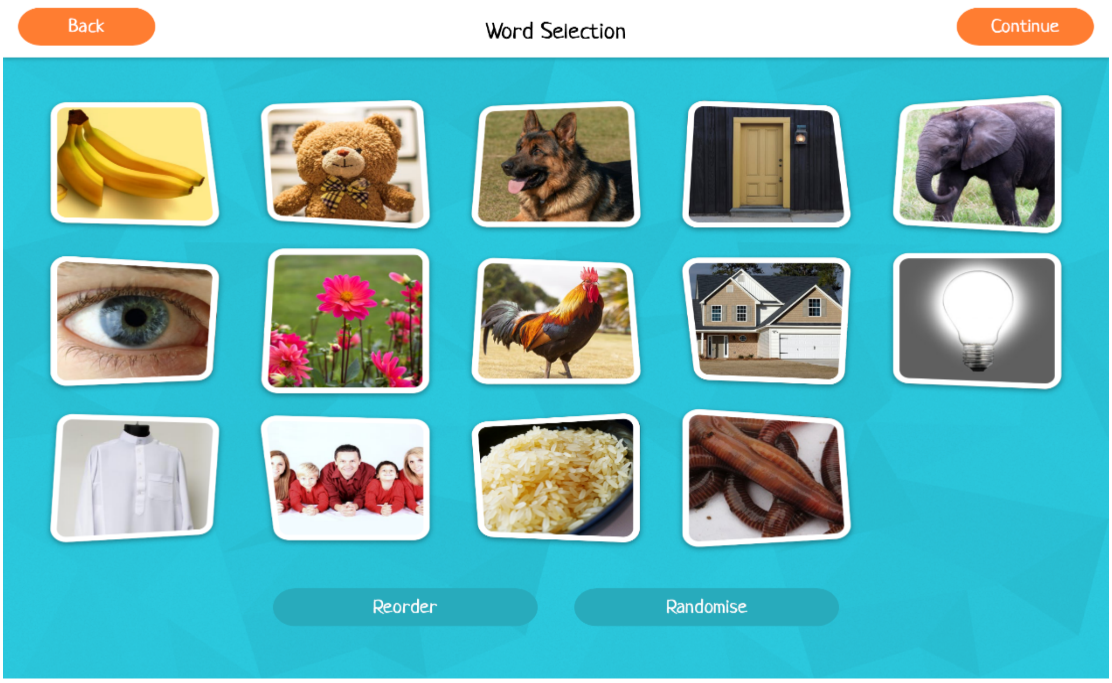
Instructions on how to use the application to test a child's hearing

Please read the following instructions and follow it carefully:

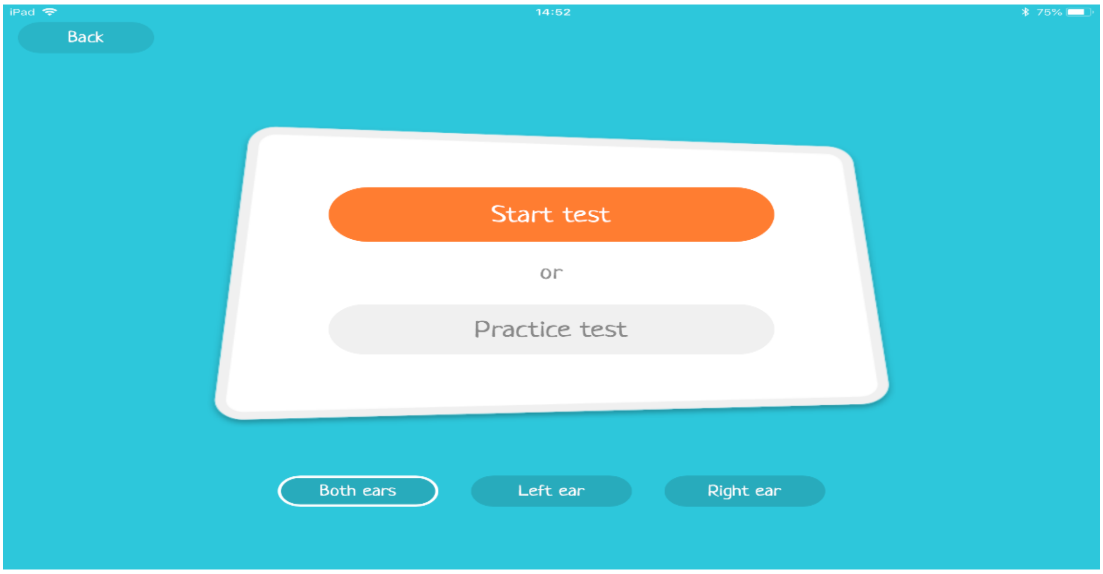
1. On the iPad, touch the PAAST icon. The following screen will be displayed. Select the language that you prefer (please note that the test will be run in Arabic language despite of the language of the instructions) and then select "Run Test ".



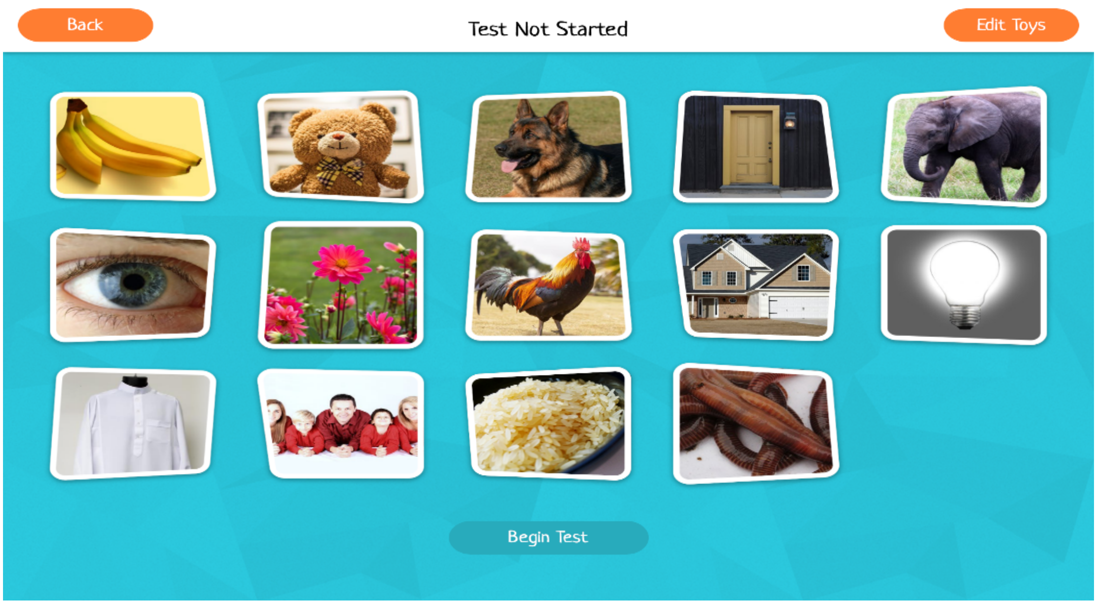
2. The following screen will appear, select "Continue".



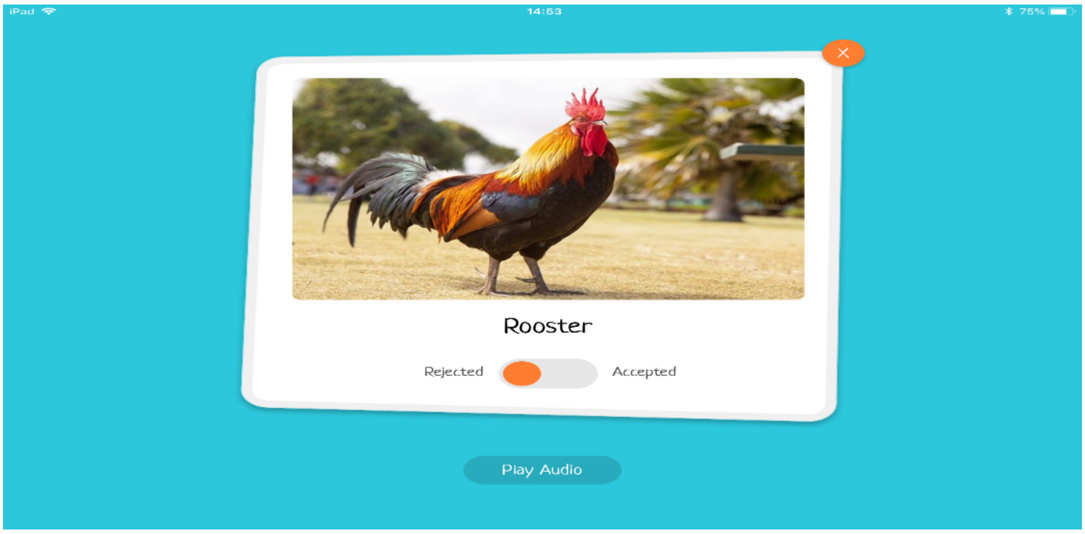
3. Although you can perform the test on each ear separately (right or left), please keep this selection as default "Both ears" for now and select "Start test".



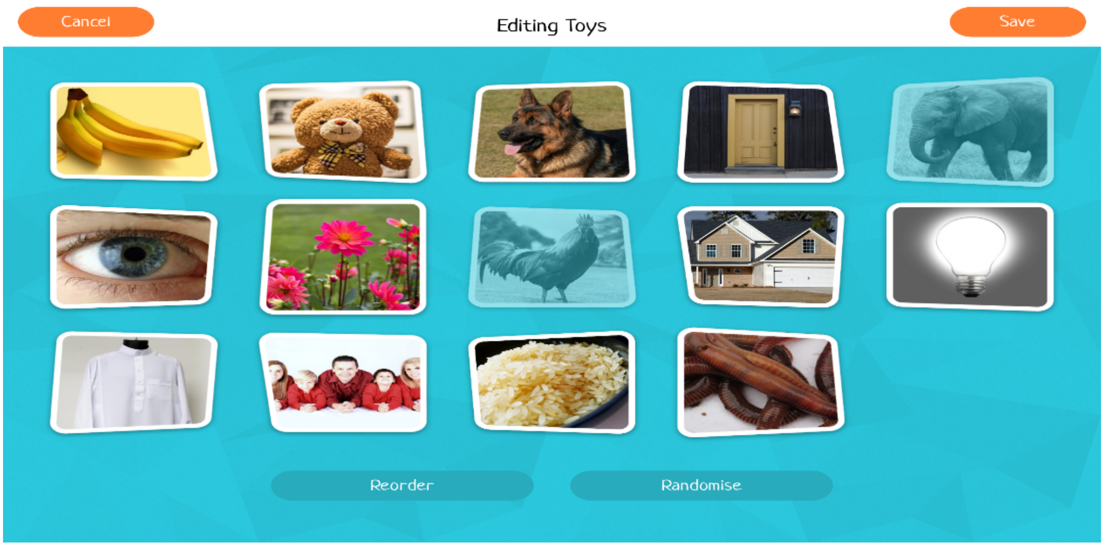
4. The following screen will appear (note that the test is not yet started). Before the test start, you will need to ensure that the child is familiar with the words. This is because we want to make sure that if the child responds incorrectly, this would be because s/he did not hear the word rather than s/he did not recognise the word. In order to do that, ask the child ("where is the "name of a picture" e.g. where is the door?) and ask the child to point to the picture you named. Ask the child about all 14 pictures that is appearing in the screen below.



If the child identifies all the words correctly, you can move on step 5.
In case the child struggled to know any of the words, select "Edit Toys" on the Right top of the screen above then touch that specific picture that represent the word the child did not recognise. The following screen will appear with the picture you selected.



Select “Rejected” and then close the window by touching the “X” symbol on the top of the picture. This action will exclude that word <u>and its pair</u> from the test as shown below (the excluded pairs are shadowed). Select “Save” to save your changes.



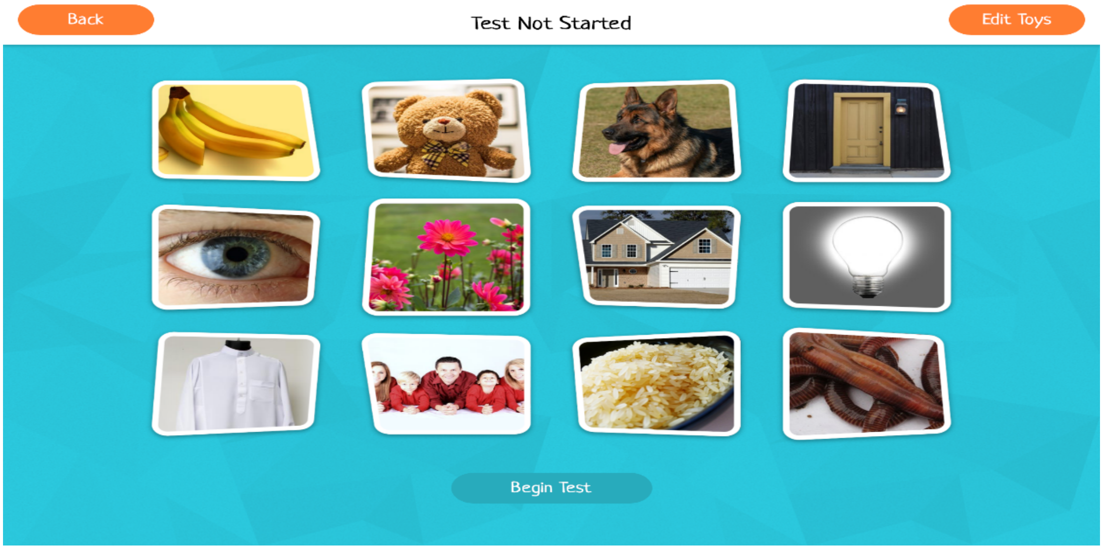
5. Before starting the test, instruct the child clearly by saying the following. “I will place earphones in your ears, you will listen to sentences just like the ones I have said to you, you will need to touch the picture of the word that you heard. Try your best to focus and listen carefully because sometimes the loudness of the word will be very low. If you are not sure what you have heard, you will need to guess.”

6. Make sure that the child has understood your instruction and have him/her seated comfortably.

7. Set the sound presentation level to a low level, listen to the sound yourself and increase it gradually until it reaches a comfortable loudness level before you place the earphone in the child's ear.

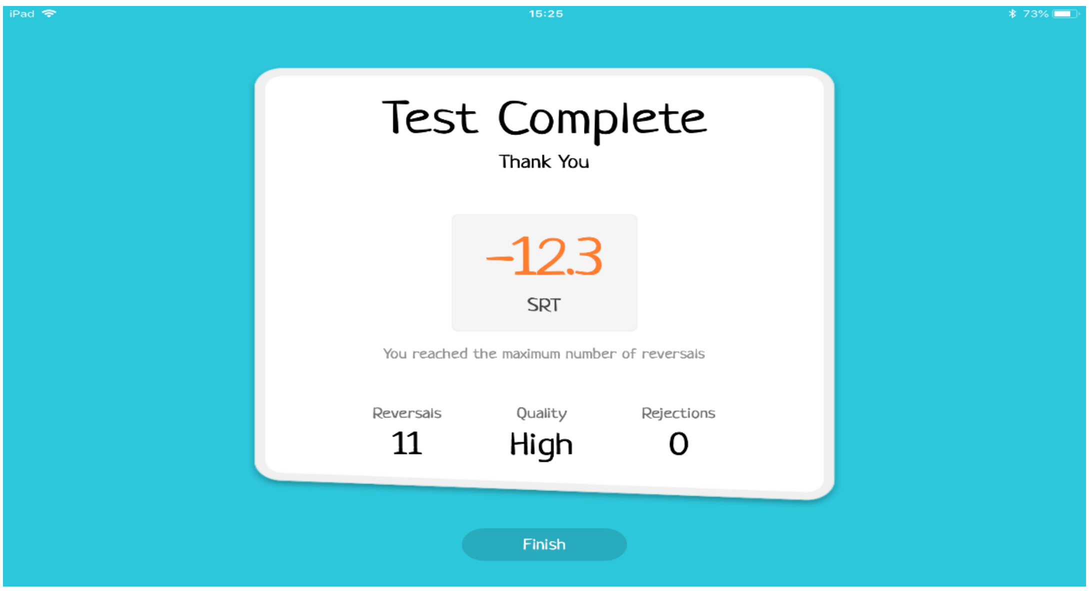
8. Place the earphones carefully in his/her both ears and make sure that it is stable.

9. On the following screen, select "Begin test". Once you select that, the test will start immediately.



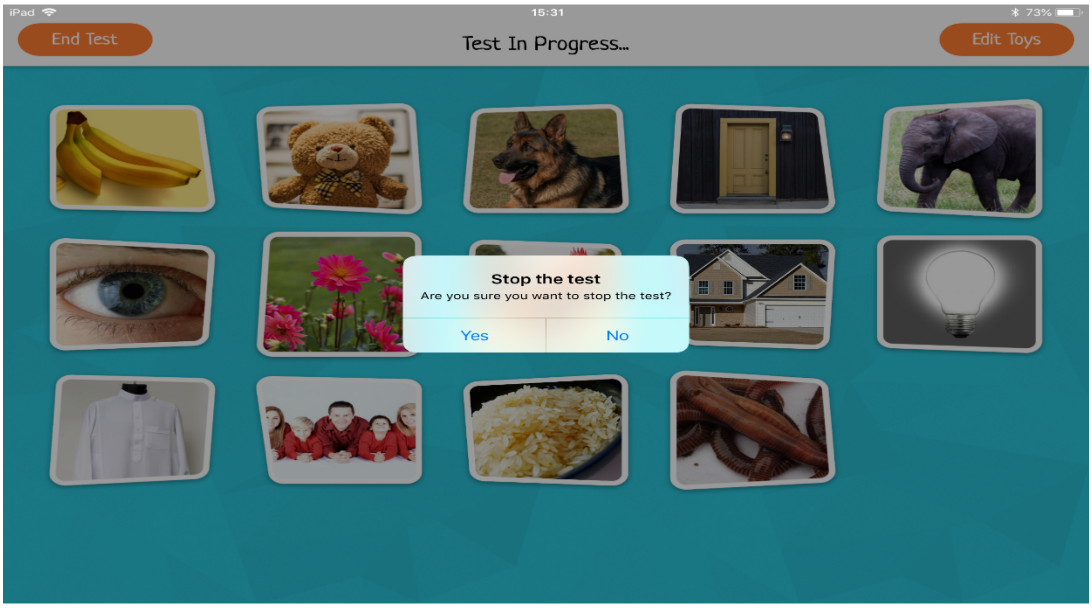
10. Wait for the test to complete (It takes about 5-6 minutes). You do not need to interfere until the test is completed. However, keep monitoring the child and make sure that s/he is attentive. If s/he says that s/he is not sure what the words is, remind her/him that s/he must guess in this case.

11. Once the test is completed, a screen with the result will display as follows, select "Finish".

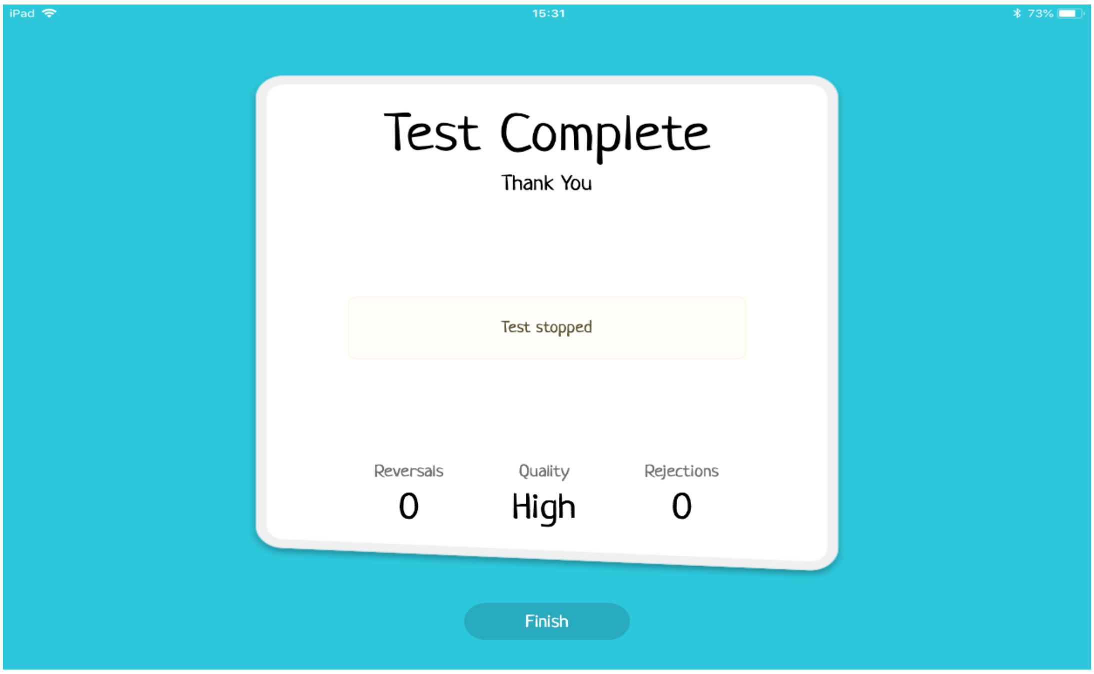


12. If you need to stop the test for any reason, select “End test”.

The following window will appear



Select “Yes”, the test will be stopped, select “Finish”.



13. Remove the earphones and thank the child for his/her cooperation.

14. Fill out the provided questionnaire.

15. If you are still to test other children, follow the same procedure.

THANK YOU ☺

Appendix H System usability questionnaire (parents)

System Usability Scale

Study title: Feasibility and Usability of the Children Auditory Picture Test (CAPT)
Researcher name: Rania Alkahtani **ERGO number:** 45826

Participant ID: ______________ Date: ___/___/___
Gender: ____________________ Age: __________
Test's Duration: __________________

Instructions: For each of the following statements, mark one box that best describes your reactions to the application *today*.

		Strongly Disagree				Strongly Agree
1.	I think that I would like to use this application frequently.	☐	☐	☐	☐	☐
2.	I found this application unnecessarily complex.	☐	☐	☐	☐	☐
3.	I thought this application was easy to use.	☐	☐	☐	☐	☐
4.	I think that I would need assistance to be able to use this application.	☐	☐	☐	☐	☐
5.	I found the various functions in this application were well integrated.	☐	☐	☐	☐	☐
6.	I thought there was too much inconsistency in this application.	☐	☐	☐	☐	☐
7.	I would imagine that most people would learn to use this application very quickly.	☐	☐	☐	☐	☐
8.	I found this application very cumbersome/awkward to use.	☐	☐	☐	☐	☐
9.	I felt very confident using this application.	☐	☐	☐	☐	☐
10.	I needed to learn a lot of things before I could get going with this application.	☐	☐	☐	☐	☐

Please provide any comments about this application:
____
____
____
____

If you found the application difficult to use or had any problem using it and are happy for us to contact you for further clarification, kindly email the researcher, Rania Alkahtani, on the following email: raa1g15@soton.ac.uk and provide her with your contact details.
This could be your phone number, your skype account, any other preferable method for her to contact you including face-to-face meeting).

Note: Please do not write your contact details on this form.

Appendix I System usability questionnaire (teachers)

System Usability Scale

Study title: Feasibility and Usability of the Children Auditory Picture Test (CAPT)
Researcher name: Rania Alkahtani **ERGO number:** 45826

Participant ID: ______________ Date: ___/___/___

Gender: ___________________ Age: __________

Number of children tested: ____________

Test Duration for one child: _________________ Test duration for all children: ____________

Instructions: For each of the following statements, mark one box that best describes your reactions to the application *today*.

		Strongly Disagree				Strongly Agree
1.	I think that I would like to use this application frequently.	☐	☐	☐	☐	☐
2.	I found this application unnecessarily complex.	☐	☐	☐	☐	☐
3.	I thought this application was easy to use.	☐	☐	☐	☐	☐
4.	I think that I would need assistance to be able to use this application.	☐	☐	☐	☐	☐
5.	I found the various functions in this application were well integrated.	☐	☐	☐	☐	☐
6.	I thought there was too much inconsistency in this application.	☐	☐	☐	☐	☐
7.	I would imagine that most people would learn to use this application very quickly.	☐	☐	☐	☐	☐
8.	I found this application very cumbersome/awkward to use.	☐	☐	☐	☐	☐
9.	I felt very confident using this application.	☐	☐	☐	☐	☐
10.	I needed to learn a lot of things before I could get going with this application.	☐	☐	☐	☐	☐

Please provide any comments about this application:

If you found the application difficult to use or had any problem using it and are happy for us to contact you for further clarification, kindly email the researcher, Rania Alkahtani, on the following email: raa1g15@soton.ac.uk and provide her with your contact details.
This could be your phone number, your skype account, any other preferable method for her to contact you including face-to-face meeting).

Note: Please do not write your contact details on this form.

List of References

Abolfotouh MA, Ghieth MM and Badawi IA (1995) Hearing loss and other ear problems among schoolboys in Abha, Saudi Arabia. *Annals of Saudi medicine* 15(4): 323-6

Absalan A, Pirasteh I, Dashti Khavidaki GA, Nasr Esfahani AA and Nilforoush MH (2013) A Prevalence Study of Hearing Loss among Primary School Children in the South East of Iran. *International journal of otolaryngology* 2013

Accenture (2017) *Survey finds growing use of digital health in Saudi Arabia*. Available from: https://www.accenture.com/sa-en/insight-saudi-consumers [Accessed 22-11-2019]

Adler PS and Winograd TA (1992) *Usability: Turning Technologies Into Tools*. Oxford: Oxford University Press, Inc.

Al'shardzhabi I and Tsygankova E (2013) [The prevalence of hearing impairment among the elementary school pupils in the city of Sana, the capital of Yemen]. *Vestnik otorinolaringologii* (2): 54-57

Al-Abduljawad KA and Zakzouk SM The prevalence of sensorineural hearing loss among Saudi children. *International Congress Series*: Elsevier 199-204

Al-Dabbous IA, Al Jam'a AH, Obeja SK, Murugan AN and Hammad HA (1996) Sensorineural hearing loss in homozygous sickle cell disease in Qatif, Saudi Arabia. *Annals of Saudi medicine* 16(6): 641-4

Al-Hanawi MK (2017) The healthcare system in Saudi Arabia: How can we best move forward with funding to protect equitable and accessible care for all? *International Journal of Healthcare* 3(2): 78-94

Al-Kandari J and Alshuaib W (2007) Newborn hearing screening in Kuwait. *Electromyography and clinical neurophysiology* 47(6): 305-313

Al-Kandari J and Alshuaib W (2009) Hearing evaluation of school children in Kuwait. *Electromyography and clinical neurophysiology* 50(6): 309-318

Al-Muhaimeed H, Al-Anazy F, Attallah M and Hamed O (2009) Cochlear implantation at King Abdulaziz University Hospital, Riyadh, Saudi Arabia: a 12-year experience. *The Journal of laryngology and otology* 123(11)

Al-Qahtani MH, Baghlab I, Chaudhary AG, Abuzenadah AM, Bamanie A, Daghistani KJ, Safieh M, Fida L and Dallol A (2010) Spectrum of GJB2 mutations in a cohort of nonsyndromic hearing loss cases from the Kingdom of Saudi Arabia. *Genetic testing and molecular biomarkers* 14(1): 79-83

Al-Rowaily MA, AlFayez AI, AlJomiey MS, AlBadr AM and Abolfotouh MA (2012) Hearing impairments among Saudi preschool children. *International Journal of Pediatric Otorhinolaryngology* 76(11): 1674-1677

Al-Shehri S (2013) *An Outlook on Future Mobile Learning in Saudi Arabia*. King Khalid University. Available from: https://www.qscience.com/docserver/fulltext/qproc/2013/3/qproc.2013.mlearn.9.pdf?expires=1579180288&id=id&accname=guest&checksum=F264828EF1E881C6776C52BFBE42A272 [Accessed 08-08-2018]

Al-Noury K (2011) Distortion product otoacoustic emission for the screening of cochlear damage in children treated with cisplatin. *The Laryngoscope* 121(5): 1081-1084

Al Khabori M and Patton MA (2008) Consanguinity and deafness in Omani children. *International journal of audiology* 47(1): 30-33

al Muhaimeed H (1996) Prevalence of sensorineural hearing loss due to toxoplasmosis in Saudi children: a hospital based study. *Internation Journal of Pediatric Otorhinolaryngology* 34(1-2): 1-8

al Muhaimeed H and Zakzouk SM (1997) Hearing loss and herpes simplex. *Journal of Tropical Pediatrics* 43(1): 20-4

Albarrak AI, Mohammed R, Almarshoud N, Almujalli L, Aljaeed R, Altuwaijiri S and Albohairy T (2019) Assessment of physician's knowledge, perception and willingness of telemedicine in Riyadh region, Saudi Arabia. *Journal of infection and public health*

Alfaqeeh G, Cook EJ, Randhawa G and Ali N (2017) Access and utilisation of primary health care services comparing urban and rural areas of Riyadh Providence, Kingdom of Saudi Arabia. *BMC health services research* 17(1): 106

Alharbi FA and Ahmed MR (2015) Evaluation of hearing among kindergarten children in Jazan (Kingdom of Saudi Arabia). *Interventional Medicine and Applied Science* 7(3): 91-94

Alkahtani R, Rowan D, Kattan N and Alwan NA (2019) Age of Identification of Sensorineural Hearing Loss and Characteristics of Affected Children: Findings from Two Cross-sectional Studies in Saudi Arabia. *International Journal of Pediatric Otorhinolaryngology* 122: 27-34

Almuayqil S, Atkins AS and Sharp B (2016) Ranking of E-Health barriers faced by Saudi Arabian citizens, healthcare professionals and IT specialists in Saudi Arabia. *Health* 8(10): 1004

Almuneef M, Memish Z, Khan Y, Kagallwala A and Alshaalan M (1998) Childhood bacterial meningitis in Saudi Arabia. *Journal of Infection* 36(2): 157-60

Alnuaimi A, Rawaf S, Hassounah S and Chehab M (2019) Use of mobile applications in the management of overweight and obesity in primary and secondary care. *The Royal Society of Medicine Journals Open* 10(3): 2054270419843826

Alotaibi YAaH, A. (2010) Comparative analysis of arabic vowels using formants and an automatic speech recognition system. *International Journal of Signal Processing, Image Processing and Pattern Recognition* 3(2): 11-20

Alradhi S, Albarrak A and Alghamdi M (2019) Acceptance of Tele-Dental Health Education Among Head and Neck Cancer Patients in Saudi Arabia. *Stud Health Technol Inform* 264: 1899-1900

Alsulame K, Khalifa M and Househ MS (2015) eHealth in Saudi Arabia: Current Trends, Challenges and Recommendations. *ICIMTH* 213: 233-236

Alusi H, Hinchcliffe R, Ingham B, Knight J and North C (1974) Arabic Speech Audiomentry. *Audiology* 13(3): 212-230

Alzannan BM (2015) The Effect of Using Ipad on the Achievement of Children in Layla Kindergarten in Saudi Arabia. *IOSR Journal of Resaerch & Method in Education* 5(1): 58-65

Amayrah MMaD, A.T. (2000) Phonetic inventories of young Arabic-speaking children. *Clinical linguistics & phonetics* 14(3): 193-215

Amayreh MM (2003) Completion of the consonant inventory of Arabic. *Journal of Speech, Language, and Hearing Research*

American Academy of Audiology (2011) *Childhood Hearing Screening Guidelines*

Available from: https://www.cdc.gov/ncbddd/hearingloss/documents/AAA_Childhood%20Hearing%20Guidelines_2011.pdf [Accessed 12-04-2016]

American Academy of Pediatrics (2007) *Year 2007 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs*. Available from: https://www.cdc.gov/ncbddd/hearingloss/documents/JCIH_2007.pdf [Accessed 10-05-2016]

American Academy of Pediatrics (2016) *Recommendations for Preventive Pediatric Health Care*. Available from: https://www.aap.org/en-us/Documents/periodicity_schedule.pdf [Accessed 12-04-2016]

American speech-Language-Hearing Association (2016) *Effects of Hearing Loss on Development*. Available from: http://www.asha.org/public/hearing/Effects-of-Hearing-Loss-on-Development/ [Accessed 11-5-2016]

American Speech-Language-Hearing Association (2019) *Childhood Hearing Screening*. Available from: https://www.asha.org/PRPSpecificTopic.aspx?folderid=8589935406&section=Key_Issues [Accessed 28-01-2020]

Arain M, Campbell MJ, Cooper CL and Lancaster GA (2010) What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC medical research methodology* 10(1): 67

Arnold L, Boyle P and Canning D (2010) Development of a paediatric audiovisual speech test in noise. *Cochlear implants international* 11(sup1): 244-248

Ashoor A (1983) Hearing levels of school children in Dammam. *The Journal of laryngology and otology* 97(1): 37-41

Ashoor A and Prochazka T (1985) Saudi Arabic speech audiometry for children. *British journal of audiology* 19(3): 229-238

Ashraf S, Parande C and Arya S (1985) Cytomegalovirus antibodies of patients in the Gizan area of Saudi Arabia. *Journal of Infectious Diseases* 152(6): 1351-1351

Baddeley A (2000) The episodic buffer: a new component of working memory? *Trends in cognitive sciences* 4(11): 417-423

Baddeley AD and Hitch G (1974) Working memory *Psychology of learning and motivation*. Elsevier 47-89

Baer T and Moore BC (1993) Effects of spectral smearing on the intelligibility of sentences in noise. *The Journal of the Acoustical Society of America* 94(3): 1229-1241

Baer T and Moore BC (1994) Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. *The Journal of the Acoustical Society of America* 95(4): 2277-2280

Bafaqeeh SA, Zakzouk SM, Al Muhaimeid H and Essa A (1994) Relevant demographic factors and hearing impairment in Saudi children: epidemiological study. *The Journal of Laryngology & Otology* 108(04): 294-298

Baltaxe C (2001) Emotional, behavioral, and other psychiatric disorders of childhood associated with communication disorders. *Handbook of early language impairment in children: Nature*: 63-125

Bamford J, Fortnum H, Bristow K, Smith J, Vamvakas G, Davies L, Taylor R, Watkin P, Fonseca S and Davis A (2007) Current practice, accuracy, effectiveness and cost-effectiveness of the school entry hearing screen. *Health Technology Assessment* 11(32): 161-168

Bangor A, Kortum P and Miller J (2009) Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4(3): 114-123

Bangor A, Kortum PT and Miller JT (2008) An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction* 24(6): 574-594

Barreira-Nielsen C and de Almeida Carneiro L (2015) Telehealth: verification of the effectiveness of an orientation program and audiological counseling for adults through teleconsultation. *Distúrbios Comun* 27(4): 696-706

Beattie RC, Barr T and Roup C (1997) Normal and hearing-impaired word recognition scores for monosyllabic words in quiet and noise. *British journal of audiology* 31(3): 153-164

Bell SA (2001) A beginner's guide to uncertainty of measurement. *National Physical Laboratory* 11

Bellman S, Mahon M and Triggs E (1996) Evaluation of the E2L Toy Test as a screening procedure in clinical practice. *British journal of audiology* 30(4): 286-296

Bener A, EIHakeem AA and Abdulhadi K (2005) Is there any association between consanguinity and hearing loss. *International Journal of Pediatric Otorhinolaryngology* 69(3): 327-333

Bernstein JG and Oxenham AJ (2006) The relationship between frequency selectivity and pitch discrimination: Sensorineural hearing loss. *The Journal of the Acoustical Society of America* 120(6): 3929-3945

Beschorner B and Hutchison A (2013) iPads as a literacy teaching tool in early childhood. *International Journal of Education in Mathematics Science and Technology* 1(1): 16-24

Bess FH, Dodd-Murphy J and Parker RA (1998) Children with minimal sensorineural hearing loss: prevalence, educational performance, and functional status. *Ear and hearing* 19(5): 339-354

Bess FH and Tharpe AM (1986a) Case history data on unilaterally hearing-impaired children. *Ear and hearing* 7(1): 14-19

Bess FH and Tharpe AM (1986b) An introduction to unilateral sensorineural hearing loss in children. *Ear and hearing* 7(1): 3-13

Bess FH, Tharpe AM and Gibler AM (1986) Auditory performance of children with unilateral sensorineural hearing loss. *Ear and hearing* 7(1): 20-26

Bitner-Glindzicz M (2002) Hereditary deafness and phenotyping in humans. *British Medical Bulletin* 63(1): 73-94

Blamey PJ, Blamey JK and Saunders E (2015) Effectiveness of a teleaudiology approach to hearing aid fitting. *Journal of Telemedicine and Telecare* 21(8): 474-8

Bland JM and Altman DG (1996) Measurement error. *BMJ: British medical journal* 313

Bland M (2015) *An Introduction to Medical Statistics* (Forth Edition). New York: Oxford University Press Inc.

Boothroyd A (1997) Auditory development of the hearing child. *Scandinavian Audiology-Supplement Only* 46: 9-16

Boothroyd A Room acoustics and speech perception. *Seminars in Hearing*: THIEME MEDICAL PUBLISHERS INC 155-166

Brand T and Kollmeier B (2002) Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. *The Journal of the Acoustical Society of America* 111(6): 2801-2810

Bright T and Pallawela D (2016) Validated Smartphone-Based Apps for Ear and Hearing Assessments: A Review. *JMIR Rehabilitation and Assistive Technology* 3(2): e13

Bristow K, Fortnum H, Fonseca S and Bamford J (2008) United Kingdom school-entry hearing screening: current practice. *Archives of Disease in Childhood* 93(3): 232-235

British Society of Audiology (2011) *Recommended Procedure: Pure-tone air-conduction and bone-conduction threshold audiometry with and without masking*. Available from: http://www.thebsa.org.uk/wp-content/uploads/2014/04/BSA_RP_PTA_FINAL_24Sept11_MinorAmend06Feb12.pdf [Accessed 30-06-2018]

Brooke J (1996) SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189(194): 4-7

Brooke J (2013) SUS: a retrospective. *Journal of usability studies* 8(2): 29-40

Brungart DS (2001) Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America* 109(3): 1101-1109

Brungart DS (2005) Informational and energetic masking effects in multitalker speech perception *Speech separation by humans and machines*. Springer 261-267

Calderon R and Naidu S (1999) Further support for the benefits of early identification and intervention for children with hearing loss. *Volta review* 100(5): 53-84

Caldwell A and Nittrouer S (2013) Speech perception in noise by children with cochlear implants. *Journal of Speech, Language, and Hearing Research* 56: 13-30

Cameron S and Dillon H (2007) Development of the listening in spatialized noise-sentences test (LISN-S). *Ear and hearing* 28(2): 196-211

Carhart R and Tillman TW (1970) Interaction of competing speech signals with hearing losses. *Archives of Otolaryngology* 91(3): 273-279

Carhart R, Tillman TW and Greetis ES (1969) Perceptual masking in multiple sound backgrounds. *The Journal of the Acoustical Society of America* 45(3): 694-703

Centers for Disease Control and Prevention (2003) Infants tested for hearing loss--United States, 1999-2001. *MMWR. Morbidity and mortality weekly report* 52(41): 981

Centers for Disease Control and Prevention (2017) *Summary of 2014 National CDC EHDI Data*. Available from: https://www.cdc.gov/ncbddd/hearingloss/2014-data/2014_ehdi_hsfs_summary_h.pdf [Accessed 03-07-2018]

Cole RA, Yan Y, Mak B, Fanty M and Bailey T The contribution of consonants versus vowels to word recognition in fluent speech. *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*: IEEE 853-856

Cone BK, Wake M, Tobin S, Poulakis Z and Rickards FW (2010) Slight-mild sensorineural hearing loss in children: audiometric, clinical, and risk factor profiles. *Ear and hearing* 31(2): 202-212

Corbin NE, Bonino AY, Buss E and Leibold LJ (2016) Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers. *Ear and hearing* 37(1): 55

Cox RM, Alexander GC and Gilmore C (1987) Development of the Connected Speech Test (CST). *Ear and hearing* 8(5 Suppl): 119S-126S

Crandell CC (1993) Speech recognition in noise by children with minimal degrees of sensorineural hearing loss. *Ear and hearing* 14(3): 210-216

Crukley J, Scollie S and Parsa V (2011) An exploration of non-quiet listening at school. *Journal of Educational Audiology* 17(1): 23-35

da Silva Costa SM, Ramos PZ, Martins FTA and Sartorato EL (2017) Genetic Diagnosis of Deafness *The Role of Pendrin in Health and Disease*. Springer 61-81

Dalzell L, Orlando M, MacDonald M, Berg A, Bradley M, Cacace A, Campbell D, DeCristofaro J, Gravel J and Greenberg E (2000) The New York State universal newborn hearing screening demonstration project: ages of hearing loss identification, hearing aid fitting, and enrollment in early intervention. *Ear and hearing* 21(2): 118-130

Daud MKM, Noor RM, Rahman NA, Sidek DS and Mohamad A (2010) The effect of mild hearing loss on academic performance in primary school children. *International Journal of Pediatric Otorhinolaryngology* 74(1): 67-70

Davis A, Reeve K, Hind S, Bamford J, Seewald R and Gravel J (2001) Children with mild and unilateral hearing impairment. *A sound foundation through early amplification* 4: 179-186

Davis AC (1989) The prevalence of hearing impairment and reported hearing disability among adults in Great Britain. *International journal of epidemiology* 18(4): 911-917

De Sousa KC, Swanepoel W, Moore DR and Smits C (2018) A Smartphone National Hearing Test: Performance and Characteristics of Users. *American journal of audiology* 27(3s): 448-454

Del Mar C, Doust J and Glasziou PP (2008) *Clinical Thinking*. BMJ Books

DeLuzio J and Girolametto L (2011) Peer interactions of preschool children with and without hearing loss. *Journal of Speech, Language, and Hearing Research* 54(4): 1197-1210

Denys S, Hofmann M, Luts H, Guérin C, Keymeulen A, Van Hoeck K, van Wieringen A, Hoppenbrouwers K and Wouters J (2018) School-age hearing screening based on speech-in-noise perception using the digit triplet test. *Ear and hearing* 39(6): 1104-1115

Dickinson DK, McCabe A, Essex MJ, Dickinson D and Neuman S (2006) A window of opportunity we must open to all: The case for preschool with high-quality support for language and literacy. *Handbook of early literacy research* 2: 11-28

Divenyi PL and Haupt KM (1997) Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. II. Correlation analysis. *Ear Hear* 18(2): 100-13

Dongre AR, Patil AP, Wahurwagh AJ, Kothari A, Burchundi K and Manohare MP (2017) Acoustical characteristics of classrooms of tropical climate. *Applied Acoustics* 121: 46-55

Dreschler WA and Plomp R (1980) Relation between psychophysical data and speech perception for hearing-impaired subjects. I. *The Journal of the Acoustical Society of America* 68(6): 1608-1615

Durlach N (2006) Auditory masking: Need for improved conceptual structure. *The Journal of the Acoustical Society of America* 120(4): 1787-1790

Durlach NI, Mason CR, Shinn-Cunningham BG, Arbogast TL, Colburn HS and Kidd Jr G (2003) Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *The Journal of the Acoustical Society of America* 114(1): 368-379

Edgerton BJ and Danhauer JL (1979) *Clinical implications of speech discrimination testing using nonsense stimuli*. University Park Press

Eisenberg LS, Shannon RV, Martinez AS, Wygonski J and Boothroyd A (2000) Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America* 107(5 Pt 1): 2704-10

El-Mahalli AA, El-Khafif SH and Al-Qahtani MF (2012) Successes and challenges in the implementation and application of telemedicine in the eastern province of Saudi Arabia. *Perspectives in health information management/AHIMA, American Health Information Management Association* 9(Fall)

El-Mouzan MI, Al-Salloum AA, Al-Herbish AS, Qurachi MM and Al-Omar AA (2007) Regional variations in the prevalence of consanguinity in Saudi Arabia. *Saudi Medical Journal* 28(12): 1881-1884

El Mouzan M, Al Salloum A, Al Herbish A, Qurachi M and Al Omar A (2008) Consanguinity and major genetic disorders in Saudi children: a community-based cross-sectional study. *Annals of Saudi medicine* 28(3): 169

Elfenbein JL, Hardin-Jones MA and Davis JM (1994) Oral communication skills of children who are hard of hearing. *Journal of Speech, Language, and Hearing Research* 37(1): 216-226

Elliott LL (1979) Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *The Journal of the Acoustical Society of America* 66(3): 651-653

Elliott LL, Connors S, Kille E, Levin S, Ball K and Katz D (1979) Children's understanding of monosyllabic nouns in quiet and in noise. *The Journal of the Acoustical Society of America* 66(1): 12-21

Eulitz C, Diesch E, Pantev C, Hampson S and Elbert T (1995) Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. *Journal of Neuroscience* 15(4): 2748-2755

Fageeh NA (2004) Prospective study of hearing loss in schools for deaf children in Assir region, Saudi Arabia. *West African journal of medicine* 22(4): 321-323

Fallon M, Trehub SE and Schneider BA (2000) Children's perception of speech in multitalker babble. *The Journal of the Acoustical Society of America* 108(6): 3023-3029

Faulkner L (2003) Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers* 35(3): 379-383

Festen J and Plomp R (1983) Relations between auditory functions in impaired hearing. *The Journal of the Acoustical Society of America* 73(2): 652-662

Field A (2013) *Discovering Statistics Using IBM SPSS Statistics* (Forth Edition). London: SAGE Publications Ltd

Field MJ and Grigsby J (2002) Telemedicine and remote patient monitoring. *Jama* 288(4): 423-425

Fletcher H (1929) *Speech and Hearing*. New York, USA: D. Van Nostrand Company, Inc.

Fletcher H (1940) Auditory patterns. *Reviews of modern physics* 12(1): 47

Fletcher H and Steinberg J (1929) Articulation testing methods. *The Bell System Technical Journal* 8(4): 806-854

Florentine M, Buus S, Scharf B and Zwicker E (1980) Frequency selectivity in normally-hearing and hearing-impaired observers. *Journal of Speech, Language, and Hearing Research* 23(3): 646-669

Fortnum H and Davis A (1997) Epidemiology of permanent childhood hearing impairment in Trent Region, 1985–1993. *British journal of audiology* 31(6): 409-446

Fortnum H, Ukoumunne OC, Hyde C, Taylor RS, Ozolins M, Errington S, Zhelev Z, Pritchard C, Benton C and Moody J (2016) A programme of studies including assessment of diagnostic accuracy of school hearing screening tests and a cost-effectiveness model of school entry hearing screening programmes. *Health Technology Assessment* 20(36)

Fortnum HM, Davis A, Summerfield AQ, Marshall DH, Davis AC, Bamford JM, Yoshinaga-Itano C and Hind S (2001) Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: questionnaire based ascertainment studyCommentary: Universal newborn hearing screening: implications for coordinating and developing services for deaf and hearing impaired children. *Bmj* 323(7312): 536

Fowler KB, Dahle AJ, Boppana SB and Pass RF (1999) Newborn hearing screening: will children with hearing loss caused by congenital cytomegalovirus infection be missed? *The Journal of pediatrics* 135(1): 60-64

Freyman RL, Balakrishnan U and Helfer KS (2004) Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America* 115(5): 2246-2256

Fulcher A, Purcell AA, Baker E and Munro N (2012) Listen up: children with early identified hearing loss achieve age-appropriate speech/language outcomes by 3years-of-age. *International Journal of Pediatric Otorhinolaryngology* 76(12): 1785-1794

Garadat SN, Abdulbaqi KJ and Haj-Tas MA (2017) The development of the University of Jordan word recognition test. *International journal of audiology* 56(6): 424-430

Gat IB and Keith RW (1978) An effect of linguistic experience: Auditory word discrimination by native and non-native speakers of English. *Audiology* 17(4): 339-345

Gelfand SA (2009) *Essentials of Audiology* (3rd Edition Edition). New York: Thieme Medical Publishers, Inc.

Gelfand SA (2010) *An Introduction to psychological and Physiological Acoustics* (Fifth Edition). London: Informa Healthcare

Gelnett D, Sumida A, Nilsson M and Soli S (1995) Development of the Hearing in Noise Test for Children (HINT-C). *Annual Meeting of the American Academy of Audiology.* Dallas, TX:

General Authority for Statistics KoSA (2015) *Annual Yearbook*. Available from: http://www.stats.gov.sa/en [Accessed 09-01-2020]

Georg-August P (2002) Unilateral sensorineural hearing impairment in childhood: analysis of 31 consecutive cases. *International journal of audiology* 41: 57-63

Gescheider GA (1997) *Psychophysics: The Fundamentals* (Third Edition). New Jersey: Lawrence Erlbaum Associates, Inc.

Giolas TG and Epstein A (1963) Comparative intelligibility of word lists and continuous discourse. *Journal of Speech and Hearing Research* 6(4): 349-358

Glasberg BR and Moore BC (1988) Psychoacoustic abilities of subjects with unilateral and bilateral cochlear hearing impairments and their relationship to the ability to understand speech. *Scandinavian audiology*

Global Media Insight (2019) *Saudi Arabia Social Media Statistics 2019*. Available from: https://www.globalmediainsight.com/blog/saudi-arabia-social-media-statistics/ [Accessed 22-11-2019]

Greenberg MT and Kusché CA (1993) *Promoting social and emotional development in deaf children: The PATHS project*. University of Washington Press

Haapaniemi J (1995) The 6 kHz acoustic dip in school-aged children in Finland. *European archives of oto-rhino-laryngology* 252(7): 391-394

Habib HS and Abdelgaffar H (2005) Neonatal hearing screening with transient evoked otoacoustic emissions in Western Saudi Arabia. *International Journal of Pediatric Otorhinolaryngology* 69(6): 839-842

Hagerman B (2002) Speech recognition threshold in slightly and fully modulated noise for hearing-impaired subjects: Umbral de reconocimiento del lenguaje para sujetos hipoacúsicos en medio de ruido parcial o completamente modulado. *International journal of audiology* 41(6): 321-329

Hall AJ, Munro KJ and Heron J (2007) Developmental changes in word recognition threshold from two to five years of age in children with different middle ear status: Cambios evolutivos en el umbral de reconocimiento del lenguaje de los dos a los cinco años en niños con una condición diferente en el oído medio. *International journal of audiology* 46(7): 355-361

Hall DMBaE, D (2003) *Health for all children* (Fourth Edition). New York: Oxford University Press Inc.

Hall JW, 3rd, Grose JH, Buss E and Dev MB (2002) Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear Hear* 23(2): 159-65

Hall SJ (2006) *The development of a new English sentence in noise test and an English number recognition test.* Unpublished Ph. D. thesis University of Southampton University

Hamill B (1988) Comparing two methods of preschool and kindergarten hearing screening. *Journal of School Health* 58(3): 95-97

Harries J and Williamson T (2000) Community-based validation of the McCormick Toy Test. *British journal of audiology* 34(5): 279-283

Harrison M and Roush J (1996) Age of suspicion, identification, and intervention for infants and young children with hearing loss: A national study. *Ear and hearing* 17(1): 55-62

Hartley DE, Wright BA, Hogan SC and Moore DR (2000) Age-related improvements in auditory backward and simultaneous masking in 6-to 10-year-old children. *Journal of Speech, Language, and Hearing Research* 43(6): 1402-1415

Helfand M, Thompson D, Davis R, McPhillips H, Homer C and Lieu T (2001) Newborn hearing screening: Systematic evidence review. *AHRQ Publication No*: 02-S001

Henderson D and Hendershott A (1991) ASL and the family system. *American Annals of the Deaf* 136(4): 325-329

Hicks CB and Tharpe AM (2002) Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research* 45(3): 573-584

Hintermair M (2000) Hearing impairment, social networks, and coping: The need for families with hearing-impaired children to relate to other parents and to hearing-impaired adults. *American Annals of the Deaf* 145(1): 41-53

Holzinger D, Weishaupt A, Fellinger P, Beitel C and Fellinger J (2016) Prevalence of 2.2 per mille of significant hearing loss at school age suggests rescreening after NHS. *International Journal of Pediatric Otorhinolaryngology*

Honeycutt AA, Grosse SD, Dunlap LJ, Schendel DE, Chen H, Brann E and al Homsi G (2003) Economic costs of mental retardation, cerebral palsy, hearing loss, and vision impairment *Using Survey Data to Study Disability: Results from the National Health Survey on Disability*. Emerald Group Publishing Limited 207-228

Hopkins K and Moore BC (2009) The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise. *The Journal of the Acoustical Society of America* 125(1): 442-446

Hopkins K, Moore BC and Stone MA (2008) Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech. *The Journal of the Acoustical Society of America* 123(2): 1140-1153

Horst JW (1987) Frequency discrimination of complex signals, frequency selectivity, and speech perception in hearing-impaired subjects. *The Journal of the Acoustical Society of America* 82(3): 874-885

Howell DC (1997) *Statistical Methods for Psychology* (Fourth Edition). United States of America: Wadsworth Publishing Company

Humaid A-HI, Abou-halawa SA, Masood KA, Nuha A-HS and Al Duways Ali Saleh AM (2014) Prevalence and risk factors of Otitis Media with effusion in school children in Qassim Region of Saudi Arabia. *International journal of health sciences* 8(4): 325

Humes LE, Allen SK and Bess FH (1980) Horizontal sound localization skills of unilaterally hearing-impaired children. *Audiology* 19(6): 508-518

Hussain T, Alghasham AA and Raza M (2011) Prevalence of hearing impairment in school children. *International journal of health sciences* 5(2 Suppl 1): 46

Hwang WJ and Jo HH (2019) Evaluation of the Effectiveness of Mobile App-Based Stress-Management Program: A Randomized Controlled Trial. *International journal of environmental research and public health* 16(21): 4270

Imtiaz F, Taibah K, Ramzan K, Bin-Khamis G, Kennedy S, Al-Mubarak B, Trabzuni D, Allam R, Al-Mostafa A and Sogaty S (2011) A comprehensive introduction to the genetic basis of non-syndromic hearing loss in the Saudi Arabian population. *BMC medical genetics* 12(1): 1

Jamieson DG, Kranjc G, Yu K and Hodgetts WE (2004) Speech intelligibility of young school-aged children in the presence of real-life classroom noise. *Journal of the American Academy of Audiology* 15(7): 508-517

Järvelin MR, Mäki–torkko E, Sorri MJ and Rantakallio PT (1997) Effect of hearing impairment on educational outcomes and employment up to the age of 25 years in northern Finland. *British journal of audiology* 31(3): 165-175

JCIH (2007) Year 2007 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. *Pediatrics* 120(4): 898-921

Jerger J, Speaks C and Trammell JL (1968) A new approach to speech audiometry. *Journal of Speech and Hearing Disorders* 33(4): 318-328

Jerger S and Jerger J (1984) *Pediatric speech intelligibility test*. Auditec of St. Louis

Jerger S, Lewis S, Hawkins J and Jerger J (1980) Pediatric speech intelligibility test. I. Generation of test materials. *International Journal of Pediatric Otorhinolaryngology* 2(3): 217-230

Johnson CE (2000) Childrens' phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research* 43(1): 144-157

Johnstone PM and Litovsky RY (2006) Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *The Journal of the Acoustical Society of America* 120(4): 2177-2189

Johnstone PM, Nábl ek AK and Robertson VS (2010) Sound localization acuity in children with unilateral hearing loss who wear a hearing aid in the impaired ear. *Journal of the American Academy of Audiology* 21(8): 522-534

Julious SA (2005) Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry* 4(4): 287-291

Kaandorp MW, De Groot AM, Festen JM, Smits C and Goverts ST (2016) The influence of lexical-access ability and vocabulary knowledge on measures of speech recognition in noise. *International journal of audiology* 55(3): 157-167

Kaandorp MW, Smits C, Merkus P, Festen JM and Goverts ST (2017) Lexical-access ability and cognitive predictors of speech recognition in noise in adult cochlear implant users. *Trends in Hearing* 21: 2331216517743887

Katz J, Chasin M, English K, Hood L and Tillery K (2015) *Handbook of Clinical Audiology* (Seventh Edition). Wolters Kluwer Health

Kebede MM and Pischke CR (2019) Popular diabetes apps and the impact of diabetes app use on self-care behaviour: A survey among the digital community of persons with diabetes on Social Media. *Frontiers in Endocrinology* 10: 135

Kenworthy O, Klee T and Tharpe AM (1990) Speech recognition ability of children with unilateral sensorineural hearing loss as a function of amplification, speech stimuli and listening condition. *Ear and hearing* 11(4): 264-270

Kidd G, Mason C, Richards V, Gallun F and Durlach N (2007) Informational masking *Auditory perception of sound sources*. New York: New York, NY: Springer 143-190

Killion MC and Niquette PA (2000) What can the pure-tone audiogram tell us about a patient's SNR loss? *The Hearing Journal* 53(3): 46-48

Kingdom F and Prins N (2010) *Psychophysics: A Practical Introduction*. London: Elsevier Limited

Kishon-Rabin L and Rosenhouse J (2000) Speech Perception Test for Arabic-Speaking Children: Prueba de percepción del habla para niños hablantes del árabe. *Audiology* 39(5): 269-277

Klien S (2001) Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & psychophysics* 63: 1421-1455

Knecht HA, Nelson PB, Whitelaw GM and Feth LL (2002) Background noise levels and reverberation times in unoccupied classrooms. *American journal of audiology*

Knight KRG, Kraemer DF and Neuwelt EA (2005) Ototoxicity in children receiving platinum chemotherapy: underestimating a commonly occurring toxicity that may influence academic and social development. *Journal of Clinical Oncology* 23(34): 8588-8596

Koopmans WJA, Goverts ST and Smits C (2018) Speech Recognition Abilities in Normal-Hearing Children 4 to 12 Years of Age in Stationary and Interrupted Noise. *Ear Hear* 39(6): 1091-1103

Kramer SE, Kapteyn TS, Festen JM and Kramer SE (1998) The self-reported handicapping effect of hearing disabilities. *Audiology* 37(5): 302-312

Kramer SE, Kapteyn TS, Festen JM and Tobi H (1996) The relationships between self-reported hearing disability and measures of auditory disability. *Audiology* 35(5): 277-287

Kumar R (2011) *Research Methodology* (Third Edition). London: SAGE Publications Ltd.

Kuriki S and Murase M (1989) Neuromagnetic study of the auditory responses in right and left hemispheres of the human brain evoked by pure tones and speech sounds. *Experimental Brain Research* 77(1): 127-134

le Clercq CMP, Labuschagne LJE, Franken MJP, Baatenburg de Jong RJ, Luijk M, Jansen PW and van der Schroeff MP (2019) Association of Slight to Mild Hearing Loss With Behavioral Problems and School Performance in Children. *JAMA Otolaryngology- Head & Neck Surgery*

Lederberg AR and Golbach T (2002) Parenting stress and social support in hearing mothers of deaf and hearing children: A longitudinal study. *Journal of Deaf Studies and Deaf Education* 7(4): 330-345

Leensen MC, de Laat JA and Dreschler WA (2011) Speech-in-noise screening tests by internet, part 1: test evaluation for noise-induced hearing loss identification. *International journal of audiology* 50(11): 823-834

Léger AC, Moore BC and Lorenzi C (2012) Abnormal speech processing in frequency regions where absolute thresholds are normal for listeners with high-frequency hearing loss. *Hearing research* 294(1-2): 95-103

Leibold LJ (2017) Speech perception in complex acoustic environments: Developmental effects. *Journal of Speech, Language, and Hearing Research* 60(10): 3001-3008

Leibold LJ and Buss E (2013) Children's identification of consonants in a speech-shaped noise or a two-talker masker. *Journal of Speech, Language, and Hearing Research*

Levitt H (1971) Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America* 49(2B): 467-477

Lewis JR (1995) IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction* 7(1): 57-78

Lewis JR (2002) Psychometric evaluation of the PSSUQ using data from five years of usability studies. *International Journal of Human-Computer Interaction* 14(3-4): 463-488

Liberman MC, Epstein MJ, Cleveland SS, Wang H and Maison SF (2016) Toward a differential diagnosis of hidden hearing loss in humans. *PLoS One* 11(9): e0162726

Lieu JE, Tye-Murray N and Fu Q (2012) Longitudinal study of children with unilateral hearing loss. *The Laryngoscope* 122(9): 2088-2095

Lin Y-Y, Chen W-T, Liao K-K, Yeh T-C, Wu Z-A and Ho L-T (2005) Hemispheric balance in coding speech and non-speech sounds in Chinese participants. *Neuroreport* 16(5): 469-473

Ling D (1978) Auditory coding and recording: An analysis of auditory training procedures for hearing impaired children. *Auditory management of hearing-impaired children*: 181-218

Litovsky RY (2005) Speech intelligibility and spatial release from masking in young children. *The Journal of the Acoustical Society of America* 117(5): 3091-3099

Lorenzi C, Gilbert G, Carn H, Garnier S and Moore BC (2006) Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences* 103(49): 18866-18869

Lovett R, Summerfield Q and Vickers D (2013) Test-retest reliability of the Toy Discrimination Test with a masker of noise or babble in children with hearing impairment. *International journal of audiology* 52(6): 377-384

Lovett RES, Kitterick PT, Huang S and Summerfield AQ (2012) The developmental trajectory of spatial listening skills in normal-hearing children. *Journal of Speech, Language, and Hearing Research*

Lü J, Huang Z, Yang T, Li Y, Mei L, Xiang M, Chai Y, Li X, Li L and Yao G (2011) Screening for delayed-onset hearing loss in preschool children who previously passed the newborn hearing screening. *International Journal of Pediatric Otorhinolaryngology* 75(8): 1045-1049

Lund AM (2001) Measuring usability with the use questionnaire12. *Usability interface* 8(2): 3-6

Lunner T (2003) Cognitive function in relation to hearing aid use. *International journal of audiology* 42: S49-S58

Lunner T and Sundewall-Thorén E (2007) Interactions between cognition, compression, and listening conditions: Effects on speech-in-noise performance in a two-channel hearing aid. *Journal of the American Academy of Audiology* 18(7): 604-617

Lyregaard P (1982) Frequency selectivity and speech intelligibility in noise. *Scandinavian Audiology. Supplementum* 15: 113-122

MacLeod A and Summerfield Q (1990) A procedure for measuring auditory and audiovisual speech-reception thresholds for sentences in noise: Rationale, evaluation, and recommendations for use. *British journal of audiology* 24(1): 29-43

MacPherson A and Akeroyd MA (2014) Variations in the slope of the psychometric functions for speech intelligibility: A systematic survey. *Trends in Hearing* 18: 2331216514537722

Madell JR (1998) *Behavioral Evaluation of Hearing in Infants and Young Children*. New York: Thieme Medical Publishers, Inc.

Madell JR and Flexer C (2008) *Pediatric Audiology: Diagnosis, Technology and Management*. New York: Thieme Medical Publishers, Inc

Maidment DW and Ferguson M (2018) An application of the Medical Research Council's guidelines for evaluating complex interventions: A usability study assessing smartphone-connected listening devices in adults with hearing loss. *American journal of audiology* 27(3S): 474-481

Maisoun AM and Zakzouk SM (2003) Hearing screening of neonates at risk. *Saudi Medical Journal* 24(1): 55-7

Martin M (1997) *Speech Audiometry* (second Edition). London, England: Whurr Publishers Ltd

Martini FH and Bartholomew EF (2003) *Essentials of Anatomy and Physiology* (Third Edition). New Jersey: Pearson Education, Inc.

Mayer C and Wells G (1996) Can the linguistic interdependence theory support a bilingual-bicultural model of literacy education for deaf students? *Journal of Deaf Studies and Deaf Education* 1(2): 93-107

McCormick B (1977) The Toy Discrimination Test: an aid for screening the hearing of children above a mental age of two years. *Public health* 91(2): 67-74

McCormick B (1988) *Screening for Hearing Impairment in Young Children*. London & Sydney: Croom Helm Ltd

McCreery RW, Spratford M, Kirby B and Brennan M (2017) Individual differences in language and working memory affect children's speech recognition in noise. *International journal of audiology* 56(5): 306-315

McCreery RW, Walker EA, Spratford M, Lewis D and Brennan M (2019) Auditory, Cognitive, and Linguistic Factors Predict Speech Recognition in Adverse Listening Conditions for Children With Hearing Loss. *Front Neurosci* 13: 1093

McCreery RW, Walker EA, Spratford M, Oleson J, Bentler R, Holte L and Roush P (2015) Speech recognition and parent ratings from auditory development questionnaires in children who are hard of hearing. *Ear and hearing* 36: 60S-75S

McElree B, Foraker S and Dyer L (2003) Memory structures that subserve sentence comprehension. *Journal of memory and language* 48(1): 67-91

McPherson B and Brouillette R (2008) *Audiology in Developing Countries*. New York: Nova Science Publishers, Inc.

Meadow-Orlans KP and Steinberg AG (1993) Effects of infant hearing loss and maternal support on mother-infant interactions at 18 months. *Journal of Applied Developmental Psychology* 14(3): 407-426

Mehra S, Eavey RD and Keamy DG (2009) The epidemiology of hearing impairment in the United States: newborns, children, and adolescents. *Otolaryngology-Head and Neck Surgery* 140(4): 461-472

Metsala JL and Walley AC (1998) Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability *Word recognition in beginning literacy*. Lawrence Erlbaum Associates Publishers 89-120

Middelweerd M, Festen J and Plomp R (1990) Difficulties with Speech Intelligibility in noise in spite of a normal pure-tone audiogram: original papers. *Audiology* 29(1): 1-7

Ministry of Health (2016) *MOH Launches the 1st Phase of Newborn Screening for Hearing-Loss and CCHD Program*. Available from: https://www.moh.gov.sa/en/Ministry/MediaCenter/News/Pages/News-2016-10-09-001.aspx [Accessed 18-07-2018]

Ministry of Health (2019a) *MOH apps for smartphones*. Available from: https://www.moh.gov.sa/en/Support/Pages/MobileApp.aspx [Accessed 22-11-2019]

Ministry of Health (2019b) *MOH: 2nd Phase of School-Based Screening Program Launched*. Available from: https://www.moh.gov.sa/en/Ministry/MediaCenter/News/Pages/News-2019-10-13-002.aspx [Accessed 15-12-2019]

Ministry of Health (2019c) *National e-health strategy*. Available from: https://www.moh.gov.sa/en/Ministry/nehs/Pages/Ehealth.aspx [Accessed 20-11-2019]

Ministry of Health (2019d) *Schoolo-Based Screening Program*. Available from: https://www.moh.gov.sa/en/Ministry/Projects/HEC/Pages/default.aspx [Accessed 15-12-2019]

Moctar EC, Riahi Z, El Hachmi H, Veten F, Meiloud G, Bonnet C, Abdelhak S, Errami M and Houmeida A (2016) Etiology and associated GJB2 mutations in Mauritanian children with non-syndromic hearing loss. *European Archives of Otorhinolaryngology*

Moeller MP (2000) Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics* 106(3): e43-e43

Moon IJ and Hong SH (2014) What is temporal fine structure and why is it important? *Korean journal of audiology* 18(1): 1

Moore B and Glasberg B (1997) A model of loudness perception applied to cochlear hearing loss. *Auditory neuroscience* 3(3): 289-311

Moore BC (2003) *An introduction to the psychology of hearing* (Fifth Edition). London, California: Academic Press

Moore BC, Glasberg BR and Hopkins K (2006) Frequency discrimination of complex tones by hearing-impaired subjects: Evidence for loss of ability to use temporal fine structure. *Hearing research* 222(1-2): 16-27

Moore DR, Whiston H, Lough M, Marsden A, Dillon H, Munro KJ and Stone MA (2019) FreeHear: A New Sound-Field Speech-in-Babble Hearing Assessment Tool. *Trends in Hearing* 23: 2331216519872378

Mukaka MM (2012) A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal* 24(3): 69-71

Myhrum M, Tvete OE, Heldahl MG, Moen I and Soli SD (2016) The Norwegian Hearing in Noise Test for Children. *Ear and hearing* 37(1): 80-92

Najmabadi H and Kahrizi K (2014) Genetics of non-syndromic hearing loss in the Middle East. *International Journal of Pediatric Otorhinolaryngology* 78(12): 2026-2036

National Health Service (2019) *The Independent Review of Adults Screening Programmes*. Available from: https://www.england.nhs.uk/wp-content/uploads/2019/02/report-of-the-independent-review-of-adult-screening-programme-in-england.pdf [Accessed 5-12-2019]

National Institute for Health Research (2012) *NIHR Evaluation, Trials and Studies Coordination Centre: Glossary*. Available from: https://www.nihr.ac.uk/glossary?letter=F&postcategory=-1 [Accessed 04-10-2019]

Ng SL, Meston CN, Scollie SD and Seewald RC (2011) Adaptation of the BKB-SIN test for use as a pediatric aided outcome measure. *Journal of the American Academy of Audiology* 22(6): 375-86

Nia J and Bance M (2001) Effects of varying unilateral conductive hearing losses on speech-in-noise discrimination: An experimental study with implications for surgical correction. *Otology & Neurotology* 22(6): 737-744

Nicholas JG (2000) Age differences in the use of informative/heuristic communicative functions in young children with and without hearing loss who are learning spoken language. *Journal of Speech, Language, and Hearing Research* 43(2): 380-394

Nielsen J (1993) *Usability Engineering*. USA: Academic Press

Nilsson M, Soli S and Gelnett D (1996) Development of the hearing in noise test for children (HINT-C). *Los Angeles, CA: House Ear Institute*

Nilsson M, Soli SD and Sullivan JA (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

Niskar AS, Kieszak SM, Holmes A, Esteban E, Rubin C and Brody DJ (1998) Prevalence of hearing loss among children 6 to 19 years of age: The Third National Health and Nutrition Examination Survey. *Jama* 279(14): 1071-1075

Niskar AS, Kieszak SM, Holmes AE, Esteban E, Rubin C and Brody DJ (2001) Estimated prevalence of noise-induced hearing threshold shifts among children 6 to 19 years of age: the Third National Health and Nutrition Examination Survey, 1988–1994, United States. *Pediatrics* 108(1): 40-43

Nittrouer S, Caldwell-Tarr A, Tarr E, Lowenstein JH, Rice C and Moberly AC (2013) Improving speech-in-noise recognition for children with hearing loss: Potential effects of language abilities, binaural summation, and head shadow. *International journal of audiology* 52(8): 513-525

Nuesse T, Steenken R, Neher T and Holube I (2018) Exploring the link between cognitive abilities and speech recognition in the elderly under different listening conditions. *Frontiers in psychology* 9

Nunez-Batalla F, Jaudenes-Casaubon C, Sequi-Canet JM, Vivanco-Allende A and Zubicaray-Ugarteche J (2016) 2014 CODEPEH recommendations: Early detection of late onset deafness, audiological diagnosis, hearing aid fitting and early intervention. *Acta Otorrinolaringologica (English Edition)* 67(1): 45-53

Olusanya B (2001) Early detection of hearing impairment in a developing country: what options? *Audiology* 40(3): 141-7

Olusanya BO (2008) Priorities for early hearing detection and intervention in sub-Saharan Africa. *International journal of audiology* 47(sup1): S3-S13

Olusanya BO, Chapchap MJ, Castillo S, Habib H, Mukari SZ, Martinez NV, Lin H-C and McPherson B (2007) Progress towards early detection services for infants with hearing loss in developing countries. *BMC health services research* 7(1): 1

Ousey J, Sheppard S, Twomey T and Palmer A (1989) The IHR-McCormick Automated Toy Discrimination test—description and initial evaluation. *British journal of audiology* 23(3): 245-249

Ozimek E, Kutzner D, Sęk A and Wicher A (2009) Development and evaluation of Polish digit triplet test for auditory screening. *Speech Communication* 51(4): 307-316

Pagarkar W, Bitner-Glindzicz M, Knight J and Sirimanna T (2006) Late postnatal onset of hearing loss due to GJB2 mutations. *International Journal of Pediatric Otorhinolaryngology* 70(6): 1119-1124

Paglialonga A, Tognola G and Grandori F (2014) A user-operated test of suprathreshold acuity in noise for adult hearing screening: the SUN (Speech Understanding in Noise) test. *Computers in biology and medicine* 52: 66-72

Palmer A, Sheppard S and Marshall D (1991) Prediction of hearing thresholds in children using an automated toy discrimination test. *British journal of audiology* 25(5): 351-356

Paludetti G, Conti G, W DIN, E DEC, Rolesi R, Picciotti PM and Fetoni AR (2012) Infant hearing loss: from diagnosis to therapy Official Report of XXI Conference of Italian Society of Pediatric Otorhinolaryngology. *Acta Otorhinolaryngologica Italica* 32(6): 347-70

Papacharalampous GX, Nikolopoulos TP, Davilis DI, Xenellis IE and Korres SG (2011) Universal newborn hearing screening, a revolutionary diagnosis of deafness: real benefits and limitations. *European archives of oto-rhino-laryngology* 268(10): 1399-1406

Park J, Hoffman L, Marquis J, Turnbull A, Poston D, Mannan H, Wang M and Nelson L (2003) Toward assessing family outcomes of service delivery: Validation of a family quality of life survey. *Journal of Intellectual Disability Research* 47(4-5): 367-384

Paus T, Zijdenbos A, Worsley K, Collins DL, Blumenthal J, Giedd JN, Rapoport JL and Evans AC (1999) Structural maturation of neural pathways in children and adolescents: in vivo study. *Science* 283(5409): 1908-1911

Pimperton H and Kennedy CR (2012) The impact of early identification of permanent childhood hearing impairment on speech and language outcomes. *Archives of Disease in Childhood*: archdischild-2011-301501

Plomp R (1978) Auditory handicap of hearing impairment and the limited benefit of hearing aids. *The Journal of the Acoustical Society of America* 63(2): 533-549

Plomp R (2002) *The Intelligent Ear: On the Nature of Sound Perception*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Plomp R and Duquesnoy A (1982) A model for the speech-reception threshold in noise without and with a hearing aid. *Scandinavian Audiology. Supplementum* 15: 95-111

Plomp R and Mimpen A (1979) Improving the reliability of testing the speech reception threshold for sentences. *Audiology* 18(1): 43-52

Potgieter J-M, Swanepoel DW, Myburgh HC, Hopper TC and Smits C (2016) Development and validation of a smartphone-based digits-in-noise hearing test in South African English. *International journal of audiology* 55(7): 405-411

Preece J (1993) *A Guide to Usability: Human Factor in Computing*. Oxford: The Open University

Priede VM and Coles R (1976) Speech discrimination tests in investigation of sensorineural hearing loss. *The Journal of Laryngology & Otology* 90(12): 1081-1092

Public Health England (2016) *NHS Newborn Hearing Screening Programme: Standards 2016-2017*. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/685452/NHSP_Standards_2016_-_17.pdf [Accessed 10-05-2016]

Rahman TTA (2018) Development and standardization of Arabic words in noise test in Egyptian children. *International Journal of Pediatric Otorhinolaryngology* 108: 1-7

Ramkumar V, John K, Selvakumar K, Vanaja C, Nagarajan R and Hall JW (2018) Cost and outcome of a community-based paediatric hearing screening programme in rural India with application of tele-audiology for follow-up diagnostic hearing assessment. *International journal of audiology* 57(6): 407-414

Ramzan K, Al-Owain M, Allam R, Berhan A, Abuharb G, Taibah K and Imtiaz F (2013) Homozygosity mapping identifies a novel GIPC3 mutation causing congenital nonsyndromic hearing loss in a Saudi family. *Gene* 521(1): 195-199

Ramzan K, Taibah K, Tahir AI, Al-Tassan N, Berhan A, Khater AM, Al-Hazzaa SA, Al-Owain M and Imtiaz F (2014) ILDR1: novel mutation and a rare cause of congenital deafness in the Saudi Arabian population. *European journal of medical genetics* 57(6): 253-258

Rance G, Barker E, Mok M, Dowell R, Rincon A and Garratt R (2007) Speech perception in noise for children with auditory neuropathy/dys-synchrony type hearing loss. *Ear and hearing* 28(3): 351-360

Rauschecker JP (1999) Auditory cortical plasticity: a comparison with other sensory systems. *Trends in neurosciences* 22(2): 74-80

Reed CM, Braida LD and Zurek PM (2009) Review of the literature on temporal resolution in listeners with cochlear hearing impairment: A critical assessment of the role of suprathreshold deficits. *Trends in amplification* 13(1): 4-43

Repovš G and Baddeley A (2006) The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience* 139(1): 5-21

Romeo G and Bittles AH (2014) Consanguinity in the contemporary world. *Human heredity* 77(1-4): 6

Rönnberg J, Lunner T, Zekveld A, Sörqvist P, Danielsson H, Lyxell B, Dahlström Ö, Signoret C, Stenfelt S and Pichora-Fuller MK (2013) The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Frontiers in systems neuroscience* 7: 31

Roschelle JM, Pea RD, Hoadley CM, Gordin DN and Means BM (2000) Changing how and what children learn in school with computer-based technologies. *The future of children* 10(2): 76-101

Ruscetta MN, Arjmand EM and Pratt SR (2005) Speech recognition abilities in noise for children with severe-to-profound unilateral hearing impairment. *International Journal of Pediatric Otorhinolaryngology* 69(6): 771-779

Sabir I and Alsaeed N (2014) A brief description of consonants in modern standard Arabic. *Linguistics and Literature Studies* 2(7): 185-189

Saiegh-Haddad E and Joshi RM (2014) *Handbook of Arabic literacy: Insights and perspectives*. Springer Science & Business Media

Sajjad M, Khattak A, Bunn J and Mackenzie I (2008) Causes of childhood deafness in Pukhtoonkhwa Province of Pakistan and the role of consanguinity. *The Journal of Laryngology & Otology* 122(10): 1057-1063

Sauro J (2011) *A practical guide to the System Usability Scale: Background, Benchmarks & Best Practice*. CreateSpace Independent Publishing Platform

Schrijver I (2004) Hereditary non-syndromic sensorineural hearing loss: transforming silence to sound. *The Journal of molecular diagnostics* 6(4): 275-284

Sekhar DL, Zalewski TR and Paul IM (2013) Variability of state school-based hearing screening protocols in the United States. *Journal of community health* 38(3): 569-574

Semeraro HD (2015) *Developing a measure of auditory fitness for duty for military personnel.* Unpublished PhD thesis University of Southampton University

Semeraro HD, Rowan D, van Besouw RM and Allsopp AA (2017) Development and evaluation of the British English coordinate response measure speech-in-noise test as an occupational hearing assessment tool. *International journal of audiology*: 1-10

Shawky RM, Elsayed SM, Zaki ME, El-Din SMN and Kamal FM (2013) Consanguinity and its relevance to clinical genetics. *Egyptian Journal of Medical Human Genetics* 14(2): 157-164

Sheikh Rashid M, Dreschler WA and de Laat J (2017) Evaluation of an internet-based speech-in-noise screening test for school-age children. *International journal of audiology* 56(12): 967-975

Shinn-Cunningham BG (2008) Object-based auditory and visual attention. *Trends in cognitive sciences* 12(5): 182-186

Shojaei E, Jafari Z and Gholami M (2016) Effect of Early Intervention on Language Development in Hearing-Impaired Children. *Iranian Journal of Otorhinolaryngology* 28(84): 13-21

Shulman S, Besculides M, Saltzman A, Ireys H, White KR and Forsman I (2010) Evaluation of the universal newborn hearing screening and intervention program. *Pediatrics* 126(Supplement 1): S19-S27

Sim J and Lewis M (2012) The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. *Journal of clinical epidemiology* 65(3): 301-308

Sinnari D, Krause P and Abulkhair M (2018) Effects of E-Games on the Development of Saudi Children with Attention Deficit Hyperactivity Disorder Cognitively, Behaviourally and Socially: An Experimental Study. Paper presented at International Conference on Universal Access in Human-Computer Interaction las Vegas, NV, USA

Smith RJ, Bale JF and White KR (2005) Sensorineural hearing loss in children. *The Lancet* 365(9462): 879-890

Smits C, Kapteyn TS and Houtgast T (2004) Development and validation of an automatic speech-in-noise screening test by telephone. *International journal of audiology* 43(1): 15-28

Smits C, Merkus P and Houtgast T (2006) How we do it: The Dutch functional hearing-screening tests by telephone and internet. *Clinical Otolaryngology* 31(5): 436-40

Smits C, Theo Goverts S and Festen JM (2013) The digits-in-noise test: assessing auditory speech recognition abilities in noise. *The Journal of the Acoustical Society of America* 133(3): 1693-1706

Smoorenburg GF (1992) Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram. *The Journal of the Acoustical Society of America* 91(1): 421-437

Spitzlei V, Herbig R and Geissler T Online hearing screening using logatomes in fluctuating noise. *Proceedings of the International Conference on Adult Hearing Screening (AHS'10)*:

Statista (2016) *Leading health appa according to consumers in Saudi Arabia in 2016, by type*. Available from: https://www.statista.com/statistics/896074/saudi-arabia-leading-health-apps-according-to-respondents-by-type/ [Accessed 22-11-2019]

Stevens G, Flaxman S, Brunskill E, Mascarenhas M, Mathers CD and Finucane M (2013) Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries. *The European Journal of Public Health* 23(1): 146-152

Stevenson J, McCann D, Watkin P, Worsfold S and Kennedy C (2010) The relationship between language development and behaviour problems in children with hearing loss. *Journal of Child Psychology and Psychiatry* 51(1): 77-83

Stika CJ, Eisenberg LS, Johnson KC, Henning SC, Colson BG, Ganguly DH and DesJardin JL (2015) Developmental outcomes of early-identified children who are hard of hearing at 12 to 18 months of age. *Early Human Development* 91(1): 47-55

Störbeck C (2012) Childhood hearing loss in the developing world. *International Journal of child health and nutrition* 1(1): 59-65

Stuart A (2005) Development of auditory temporal resolution in school-age children revealed by word recognition in continuous and interrupted noise. *Ear and hearing* 26(1): 78-88

Stuart A (2008) Reception thresholds for sentences in quiet, continuous noise, and interrupted noise in school-age children. *Journal of the American Academy of Audiology* 19(2): 135-146

Stuart A, Givens GD, Walker LJ and Elangovan S (2006) Auditory temporal resolution in normal-hearing preschool children revealed by word recognition in continuous and interrupted noise. *The Journal of the Acoustical Society of America* 119(4): 1946-1949

Sullivan JR, Osman H and Schafer EC (2015) The effect of noise on the relationship between auditory working memory and comprehension in school-age children. *Journal of Speech, Language, and Hearing Research* 58(3): 1043-1051

Summerfield Q, Palmer A, Foster J, Marshall D and Twomey T (1994) Clinical evaluation and test-retest reliability of the IHR-McCormick Automated Toy Discrimination Test. *British journal of audiology* 28(3): 165-179

Tabri D, Abou Chacra KM and Pring T (2011) Speech perception in noise by monolingual, bilingual and trilingual listeners. *International Journal of Language and Communication Disorders* 46(4): 411-22

Taha AA, Pratt SR, Farahat TM, Abdel-Rasoul GM, Albtanony MA, Elrashiedy A-LE, Alwakeel HR and Zein A (2010) Prevalence and risk factors of hearing impairment among primary-school children in Shebin El-kom District, Egypt. *American journal of audiology* 19(1): 46-60

Talarico M, Abdilla G, Aliferis M, Balazic I, Giaprakis I, Stefanakis T, Foenander K, Grayden DB and Paolini AG (2007) Effect of age and cognition on childhood speech in noise perception abilities. *Audiology and Neurotology* 12(1): 13-19

Taylor B (2003) Speech-in-noise tests: How and why to include them in your basic test battery. *The Hearing Journal* 56(1): 40-42

Tickle-Degnen L (2013) Nuts and bolts of conducting feasibility studies. *American Journal of Occupational Therapy* 67(2): 171-176

Tiitinen H, Sivonen P, Alku P, Virtanen J and Näätänen R (1999) Electromagnetic recordings reveal latency differences in speech and tone processing in humans. *Cognitive Brain Research* 8(3): 355-363

Tullis TS and Stetson JN (2004) A Comparison of Questionnaires for Assessing Website Usability. Paper presented at Usability Professionals Association (UPA) 2004 Conference Minneapolis

Turkestani MH (2015) The Effect of iPad on School Preparedness among Preschool Children with Hearing-Impairments. *International Education Studies* 8(11): 5-62

Updike CD (1994) Comparison of FM auditory trainers, CROS aids, and personal amplification in unilaterally hearing impaired children. *Journal of the American Academy of Audiology* 5: 204-204

Uus K and Bamford J (2006) Effectiveness of population-based newborn hearing screening in England: ages of interventions and profile of cases. *Pediatrics* 117(5): e887-e893

Vaez N, Desgualdo-Pereira L and Paglialonga A (2014) Development of a test of suprathreshold acuity in noise in Brazilian Portuguese: a new method for hearing screening and surveillance. *BioMed research international* 2014

Vaillancourt V, Laroche C, Giguere C and Soli SD (2008) Establishment of age-specific normative data for the Canadian French version of the hearing in noise test for children. *Ear and hearing* 29(3): 453-466

Van Naarden K, Decouflé P and Caldwell K (1999) Prevalence and characteristics of children with serious hearing impairment in metropolitan Atlanta, 1991–1993. *Pediatrics* 103(3): 570-575

Van Wijngaarden SJ, Steeneken HJ and Houtgast T (2002) Quantifying the intelligibility of speech in noise for non-native listeners. *The Journal of the Acoustical Society of America* 111(4): 1906-1916

Vickers DA, Moore BC, Majeed A, Stephenson N, Alferaih H, Baer T and Marriage JE (2018) Closed-Set Speech Discrimination Tests for Assessing Young Children. *Ear and hearing* 39(1): 32-41

Vojnovic M and Subotic M QuickSIN test for children. *19thTelecommunications Forum (TELFOR) Proceedings of Papers*:

Wake M, Poulakis Z, Hughes E, Carey-Sargeant C and Rickards F (2005) Hearing impairment: a population study of age at diagnosis, severity, and language outcomes at 7–8 years. *Archives of Disease in Childhood* 90(3): 238-244

Watkin P, Baldwin M and Laoide S (1990) Parental suspicion and identification of hearing impairment. *Archives of Disease in Childhood* 65(8): 846-850

Watson CS, Kidd GR, Miller JD, Smits C and Humes LE (2012) Telephone screening tests for functionally impaired hearing: Current use in seven countries and development of a US version. *Journal of the American Academy of Audiology* 23(10): 757-767

Weisel A, Most T and Efron C (2005) Initiations of social interactions by young hearing impaired preschoolers. *Journal of Deaf Studies and Deaf Education* 10(2): 161-170

Wetherill G and Levitt H (1965) Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology* 18(1): 1-10

Whyte MB and Kelly P (2018) The normal range: it is not normal and it is not a range. *Postgraduate medical journal* 94(1117): 613-616

Wichmann FA and Hill NJ (2001) The psychometric function: I. Fitting, sampling, and goodness of fit. *Attention, Perception, & Psychophysics* 63(8): 1293-1313

Wightman F and Allen P (1992) Individual differences in auditory capability among preschool children *Developmental psychoacoustics* American Psychological Association 113–133

Wilson RH (2003) Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. *Journal of the American Academy of Audiology* 14(9): 453-470

Wilson RH, Farmer NM, Gandhi A, Shelburne E and Weaver J (2010) Normative data for the Words-in-Noise Test for 6-to 12-year-old children. *Journal of Speech, Language, and Hearing Research*

Wilson RH and McArdle R (2005) Speech signals used to evaluate functional status of the auditory system. *Journal of Rehabilitation Research & Development* 42

Wilson RH and Weakley DG The use of digit triplets to evaluate word-recognition abilities in multitalker babble. *Seminars in Hearing*: Published in 2004 by Thieme Medical Publishers, Inc., 333 Seventh Avenue ... 93-111

Wolff R, Hommerich J, Riemsma R, Antes G, Lange S and Kleijnen J (2010) Hearing screening in newborns: systematic review of accuracy, effectiveness, and effects of interventions after screening. *Archives of Disease in Childhood* 95(2): 130-135

Wong LL, Chen Y and Leung KP (2019) The Cantonese Hearing in Noise Test for Children. *Trends in Hearing* 23: 2331216519837128

Wong S (2015) Mobile Digital Devices and Preschoolers' Home Multiliteracy Practices. *Language and Literacy* 17(2): 75-90

Wood SA, Sutton GJ and Davis AC (2015) Performance and characteristics of the Newborn Hearing Screening Programme in England: The first seven years. *International journal of audiology* 54(6): 353-358

World Health Organization (2006) *Deafness and hearing impairment: Fact sheet No. 300*. Available from: http://www.who.int/mediacentre/factsheets/fs300/en/]

World Health Organization (2019a) *Deafness and hearing loss*. Available from: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss [Accessed 20-11-2019]

World Health Organization (2019b) *Deafness prevention*. Available from: https://www.who.int/deafness/hearWHO/en/ [Accessed 22-11-2019]

World Health Organization (2019c) *eHealth*. Available from: https://www.who.int/ehealth/en/ [Accessed 20-11-2019]

Wu W, Lü J, Li Y, Kam ACS, Tong MCF, Huang Z and Wu H (2014) A new hearing screening system for preschool children. *International Journal of Pediatric Otorhinolaryngology* 78(2): 290-295

Xu L, Thompson CS and Pfingst BE (2005) Relative contributions of spectral and temporal cues for phoneme recognition. *The Journal of the Acoustical Society of America* 117(5): 3255-3267

Yang H-M, Hsieh Y-J and Wu J-L (2012) Speech recognition performance under noisy conditions of children with hearing loss. *Clinical and experimental otorhinolaryngology* 5(Suppl 1): S73-S75

Yasui N, Adachi N, Kato M, Koh K, Asanuma S, Sakata H and Hanada R (2014) Cisplatin-induced hearing loss: the need for a long-term evaluating system. *Journal of pediatric hematology/oncology* 36(4): e241-e245

Yoshinaga-Itano C, Sedey AL, Coulter DK and Mehl AL (1998) Language of early-and later-identified children with hearing loss. *Pediatrics* 102(5): 1161-1171

Yuen KCP, Qiu XY, Mou HY and Xi X (2019) The MAndarin spoken word-Picture IDentification test in noise-Adaptive (MAPID-A) measures subtle speech-recognition-in-noise changes and spatial release from masking in very young children. *PLoS One* 14(1): e0209768

Zakzouk S, El-Sayed Y and Bafaqeeh SA (1993) Consanguinity and hereditary hearing impairment among Saudi population. *Annals of Saudi medicine* 13(5): 447-50

Zakzouk SM and al-Muhaimeed HS (1996) Hearing impairment among "at risk' children. *International Journal of Pediatric Otorhinolaryngology* 34(1-2): 75-85

Zakzouk SM, Fadle KA and al Anazy FH (1995) Familial hereditary progressive sensorineural hearing loss among Saudi population. *International Journal of Pediatric Otorhinolaryngology* 32(3): 247-55

Zakzouk SM and Hossain A (1994) Hearing impairment among children in Saudi Arabia: familial incidence and potential risk factors. *International Journal of Pediatric Otorhinolaryngology* 29(2): 111-120

Zannin PHT and Zwirtes DPZ (2009) Evaluation of the acoustic performance of classrooms in public schools. *Applied Acoustics* 70(4): 626-635

Zhang C, Arnott SR, Rabaglia C, Avivi-Reich M, Qi J, Wu X, Li L and Schneider BA (2016) Attentional modulation of informational masking on early cortical representations of speech signals. *Hearing research* 331: 119-130

Zheng Y, Soli SD, Wang K, Meng J, Meng Z, Xu K and Tao Y (2009) Development of the Mandarin pediatric speech intelligibility (MPSI) test. *International journal of audiology* 48(10): 718-728

Zokoll MA, Hochmuth S, Warzybok A, Wagener KC, Buschermöhle M and Kollmeier B (2013) Speech-in-noise tests for multilingual hearing screening and diagnostics1. *American journal of audiology*

Zokoll MA, Wagener KC, Brand T, Buschermöhle M and Kollmeier B (2012) Internationally comparable screening tests for listening in noise in several European languages: The German digit triplet test as an optimization prototype. *International journal of audiology* 51(9): 697-707