UNIVERSITY OF
EASTERN FINLAND

Listening Effort in Mainstream Early Language Education of Hard-of-Hearing Students

Can Buldu
Early Language Education for
Intercultural Communication
University of Eastern Finland
Philosophical Faculty
School of Applied Educational
Science and Teacher Education
30.05.2024

University of Eastern Finland, Philosophical Faculty
School of Applied Educational Science and Teacher Education
Master’s Degree Program in Early Language Education for Intercultural Communication
Buldu Can: Listening Effort in Mainstream Early Language Education of Hard-of-Hearing Students
Thesis, 77 pages, 10 pages of appendices
Thesis Supervisors: Prof. Ritva Kantelinen, Prof. Aarno Dietz
Mentor: Dr. Hana Vrzáková
May 2024

ABSTRACT

With the advent of cochlear implants and hearing aids, a significant number of hard-of-hearing children are joining their normal hearing peers in mainstream classrooms and take foreign language courses. While studies on foreign language competence show that cochlear-implanted students can perform on par with their normal-hearing peers, they often exhibit lower performance in listening tasks, likely due to increased listening effort. A phenomenon called "listening effort" might be a factor. Listening effort is a concept of cognitive exertion during a listening task. Even with the help of hearing aids and cochlear implants, hard-of-hearing students experience more listening effort than normal standards. In this study, a mixed method approach is adopted to explore the phenomenon from a wider perspective.

A literature review is conducted to reach a cumulative summary of the previous works documenting hard-of-hearing students in mainstream education and account of the cognitive load theory which contextualizes listening effort. Then, a group of five adult participants took part in a hearing test to simulate the adverse listening conditions experienced by hard-of-hearing students. They complete listening tasks in easy and hard settings, during which their self-reported effort scores and task performances are compared. The results of the test suggest that while task performance does not significantly vary in increasingly difficult conditions, the listener experiences more discomfort associated with listening effort. Subsequently, the literature review finds that hard-of-hearing children encounter various challenges in mainstream education, but academic success is achieved through considerations of early intervention and educational support. The quantitative and qualitative findings are presented in complementation to each other, and the study finds that hard-of-hearing students may experience more listening effort in complex tasks and multitasking activities, limiting the effectiveness of instruction. Measures taken to effectively differentiate instruction and provide support can provide the environment necessary for academic success.

Keywords: Hard-of-hearing students, cochlear implants, mainstreaming, early language education, listening effort, cognitive load.

Contents

Tables and Figures

List of Tables

List of Figures

Abbreviations

CI: Cochlear implant.

HA: Hearing aid.

NH: Normal hearing.

DHH: Deaf or hard-of-hearing.

LE: Listening effort.

CLT: Cognitive load theory.

TEPR: Task-evoked pupillary response.

RT: Response time.

Terms and Definitions

Hard-of-hearing: A person who experience mild to severe hearing impairment.

Cochlear Implant: A hearing device inserted through the inner ear via surgery, that transforms soundwaves into electrical impulses and directly stimulates the hearing nerves in the cochlea.

Hearing Aid: A hearing device that amplifies the soundwaves and releases them back into the ear via a speaker.

Mainstreaming: The act of accommodating traditionally separately educated groups of students into mainstream education.

Listening Effort: The allocation of mental resources for the completion of a listening task, and the experienced sense of difficulty.

Pupillometry: The measurement of various parameters of the pupil size and position, including “TEPR” (task-evoked pupil response), which is the pupil dilation reactive to a cognitive task.

Dual task paradigm: The test paradigm in which the participant completes two simultaneous or sequential tasks. A dual task comprised of a main cognitive task and a secondary reaction task is often used in listening effort and cognitive load studies.

1. Introduction

In the beginning of my career as a teacher, I taught English to a classroom of 20 students, including a hard-of-hearing student using a hearing aid. Inspired by this student's joy and success, I became interested in teaching foreign languages to children with hearing difficulties — a niche topic in early language education. In his 2013 essay, Kontra presents the accounts of hard-of-hearing Hungarian adults who recall their experiences of learning English at school. Kontra (2013) notes that a few decades earlier, there was a lack of support both in terms of administrative resources and teacher training. As the topic remains a niche within the larger field of early language education, there are grounds to prompt a call for extended research and examination of early language education for the hard-of-hearing.

Today, as Pritchard (2013, p.115) also attests, 90% of children with hearing impairments benefit from the early diagnosis and therapy made possible by modern medicine, as well as the use of hearing aids and cochlear implants (CI) which elevate their hearing performance to be able to attend the same foreign language courses as their peers in local schools. Indeed, Wong (2005) reports the L2 learning outcomes of three cochlear implanted students who showed considerable progress in their language proficiency. Yet, the same report (Wong, 2005) suggests that although the cochlear implanted (CI) children can learn a second language, their language development throughout their childhood shows a lower incremented pace in comparison with their normal hearing (NH) peers. This delay in language development is dependent on the timing of interventive measures, and studies show early implantation produces age-appropriate language development (Cuda et al., 2014; Shojaei et al., 2016; Ruben, 2018)

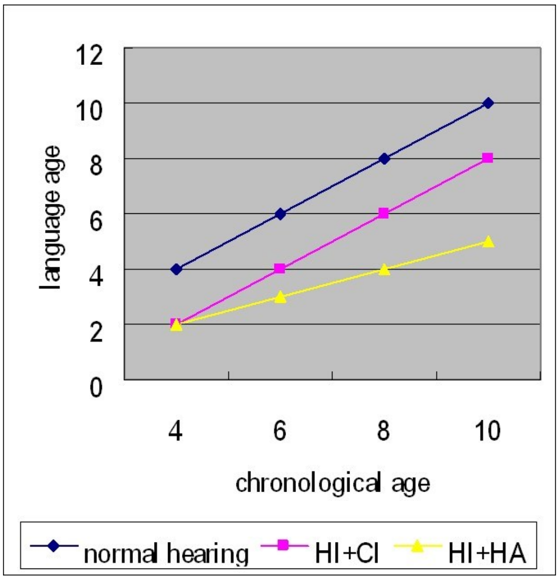

Figure 1: The figure from Wong (2005) describes the language development of children who are of normal hearing, cochlear implanted (HI + CI) and using hearing aids (HI + HA).

On the same topic, the 2023 study of Everhardt et al. displays an interesting finding. In the study, the English language proficiency progress of a group of Dutch CI students are compared with that of their NH peers. The results suggest that while the orthographic competence of the two groups is on par, the CI users' scores suffer in listening tasks. The study is quite recent, and the control group were adolescents who had been fitted with modern cochlear implants at an early age (Everhardt et al., 2023). In addition, through many studies on the field (Sharma et al., 2020; Lonka, 2014, p.19-23) it can be said that early cochlear implantation can allow language competency to develop adequately throughout childhood. While it is an anticipated result that the CI group could score lower in listening tests, what factors might influence this difference in the presence of CIs which complement hearing? To this question, "listening effort" might be one of the answers.

In their literature review, McGarrigle et al. (2014) propose the following definition for listening effort: *"...the mental exertion required to attend to, and understand, an auditory message*.". In contrast to a concept such as hearing accuracy or hearing performance, listening effort (LE) is more closely related to how much mental resources are used to accomplish a listening task (McGarrigle et al., 2014). It is a measurement of the level of attention and cognitive focus required for a person to make sense of a heard stimulus. By these definitions, it is possible to think that while a student with CIs could hear an audio recording as well as a NH student, it might take more of their attention to understand and work with the information.

In Shields et al.'s (2022) article, it is asserted that children with hearing impairments experience higher levels of stress and occasions of burnout compared to NH children, and they are one of the most affected groups from heightened levels of listening effort. Hakuta et al. (2003) and Singleton (2001) both display results that a young age of initiating contact with learning a second language is a contributing factor in determining the future proficiency, which magnifies the effects of hearing impairment and listening effort in children's early language education. Due to these concerns, demand is placed for more research on the effects of listening effort on early language education of hard-of-hearing students. As for previous literature, most of the similar studies on this topic focuses on the second language acquisition of children with CIs (Soleymani et al., 2016; Waltzman et al., 2003), being concerned more on learning a language due to immigration and less on a language classroom context, while making no references to listening effort. To fill this gap, this study aims to explore the effects of listening effort on listening tasks that might be encountered in a foreign language classroom, and how listening effort affects the early language education of hard-of-hearing children.

For this exploration, the study follows a mixed method (Schoonenboom & Johnson, 2017) approach. Both quantitative and qualitative methods are used to reach a wider understanding of the phenomenon, For the quantitative step, a listening test is performed with 5 adult participants of normal hearing, with guidance and equipment provided from the Kuopio University Hospital and University of Eastern Finland. The listening test is set to simulate easy and hard hearing conditions, and how perceived listening effort could affect performance during a listening task. For the qualitative step, a literature review is conducted on the topics of cognitive load theory and education of hard-of-hearing students to explore their experiences in mainstream early language education and contextualize the effect of listening effort. The data collection and analysis methods of both quantitative and qualitative methods will intertwine and support the coverage of each other to provide a better understanding of the phenomenon.

The research paper is organized in 7 sections, with two additional sections for appendices and references. The following section *2. Background and Concepts* will introduce the background and concepts central to the study and provide definitions of recurring terms. Section *3. Aims of the Study and Research Questions* will present the research questions, the aims of the study and the expected results. Section *4. Methods* will provide a description of the methods used throughout the study, beginning with the design of the material, followed by an account of the collection of data. A consideration of research ethics, validity and reliability is also included in section 4. Section 5. *Analysis* will demonstrate the analysis of data acquired from quantitative and qualitative methods. Section *5. Results and Discussion* will present a review of the findings from data analysis, and a discussion of the results. The final section *6. Conclusion* will provide a summary by forming a synthesis of quantitative and qualitative data in regard to the research questions and will present conclusions that can be drawn, alongside recommendations for future research. The ethical forms and a summarised handbook article of helpful classroom practices for teachers of hard-of-hearing students can be found in *Appendices*.

2. Background and Concepts

2.1. Hearing Loss

Before advancing into the more complex components of the study, an overview of the core concept of hearing impairment will be presented to establish clarity. According to the definition provided by Disabilities, Opportunities, Internetworking, and Technology (DO-IT) Center of the University of Washington, hearing impairment refers to any extent of difficulty with hearing experienced by people (DO-IT, 2022). Similarly, WHO (World Health Organization, 2018) provides that any person with a reduced ability to hear in comparison to someone with normal hearing is said to experience hearing loss. As referenced by DO-IT (2022), the term "hard-of-hearing" will be used in this study to define the population of people who experience any degree of hearing loss and resulting detriments to hearing.

The extent to which people hear is often described with sound levels measured in decibels (dB). A conversation in a low volume of sound is described as usually resting at a level of 30 dB, a normal conversation at 60 dB and loud sounds which might potentially damage hearing at above 70 dB (Riva, 2023). A person not being able to hear below the threshold of 20dB imply hearing impairment, and a hearing threshold of above 35 dB is defined as being disabling (World Health Organization, 2024). The impairment might be in mild levels, as well as being profound to a degree of highly affecting the quality of life and requiring medical attention, such as the implementation of hearing aids and cochlear implants (DO-IT, 2022). Felman (2023) offers in his article that deafness, a state associated with not being able to hear sounds below 70 dB, describes the state a person reaches at the point of not being able to comprehend spoken discourse, even in the presence of increased volume of sound. World Health Organisation (2024) offers that among other measures to prevent hearing loss through public health measures, early diagnosis and intervention holds importance for rehabilitation of hearing. Implantation is listed chiefly among the intervention methods for mild to severe cases. The main population that this study is concerned with is hard-of-hearing children who have benefited from the means of such methods and have been included in mainstream education.

2.2. Hearing Aids and Cochlear Implants

As mentioned before, most children who suffer from moderate to profound hearing loss find that they can attend their local schools and follow mainstream education, same as their peers (Pritchard, 2013, p.115). What makes this development possible is the early intervention and usage of hearing aids and cochlear implants. While both devices serve the same purpose of ameliorating the lives of people who experience hearing impairment, they hold their differences both in terms of operating principles and effects on the language development of a young user.

Hearing aids are electronic devices which use a microphone component to amplify sound and use speakers to release it back into the ear (U.S. Department of Health and Human Services, 2022). The working principle is to make sounds from the environment louder and rely on the surviving hair cells to facilitate hearing. The main advantage of hearing aids is the accessibility and practicality, with over-the-counter models readily accessible without the need for medical intervention. The devices are placed either behind or within the ear and can be used in one or both ears. Hearing aids rely on surviving hair cells and can be ineffective in moderate to severe hearing loss cases where the inner ear is damaged (U.S. Department of Health and Human Services, 2022). Typically utilised in mild to moderate hearing impairment scenarios, people with more severe forms of hearing loss are often referred to the usage of cochlear implants.

Cochlear implants (CIs) are typically used by hard-of-hearing people who have moderately or severely impaired hearing. They are similar in purpose to hearing aids in that they complement hearing, although the method of working is different. Rather than amplifying the remaining ability to hear, CIs bypass the middle ear and parts of the inner ear, and directly stimulate the hearing nerves in the cochlea (U.S. Department of Health and Human Services, 2024). CIs collect the sound from the environment with a microphone component, which is then converted into electrical impulses that are transmitted to the auditory nerve. The implants are inserted into the ear with a surgical procedure, after which patients often describe a “robotic” or “mechanical” sensation of sound (Friesen, 2019). As CIs introduce a new way of sensing sound directly through the representation of sounds as electrical impulses to the auditory nerve, a period of adjustment and learning is needed (U.S. Department of Health and Human Services, 2024). The patients receive therapy and exercise their hearing during this time, which can take between 3 to 6 months or more (Cleveland Clinic, 2023). It is important to note that neither hearing aids or CIs restore natural hearing, instead amplifying existing hearing or providing a different representation of sound respectively.

Since 2020, CIs are referred for implantation to children with hearing loss aged over 9 months in USA. (U.S. Department of Health and Human Services, 2024). Shojaei et al.’s (2016) study reports that children are implanted between 6 to 12 months of life in Iran with beneficial results. Cochlear implantation and therapy are recommended for children with severe hearing loss before 18 months of age (U.S. Department of Health and Human Services, 2024). It is demonstrated in relevant studies that early implantation for children result in development of language skills appropriate to their age (Ruben, 2018). Other studies (Sharma et al., 2020; Lonka, 2014, p.19-23; Geers et al., 2003) report that early implanted children’s linguistic development is on par with their normal hearing peers, which also translates to successful inclusion in mainstream education. The effect of CI in children’s language development is further displayed in contrast to hearing aid usage, as Torre et al.’s (2012) review suggests CI children show considerably higher linguistic progress compared to HA children.

2.3. Mainstream Education for Hard-of-Hearing Students

In his 2013 essay, Pritchard (2013, p.116) informs that the British Sign Language was being taught in Norway to students with hearing impairments as a foreign language starting from 1997 onwards, although with the emergence of hearing complementary technology, an increasing portion of hard-of-hearing students now participate in mainstream education. This is not without its nuances, and due to the nature of these devices a considerable deal of variance exists among the population of students in question (Pritchard, 2013, p.116). The goal of providing early language education to hard-of-hearing students is now encompassed within the methods of including them in mainstream education; a multidimensional task that involves elements of teacher training, educational methodology and educational planning.

The most important factor for implementing the required adjustments to the classroom to facilitate learning for the hard-of-hearing students seems to be to equip the teacher with the required knowledge and skills. It is emphasized that the primary item in the toolkit of a teacher of hard-of-hearing students should be knowledge of the technological equipment (Archbold & Mayer, 2012; Nussbaum, 2018). Blair & Langan's study (2000) shows that teachers having the knowledge to fix basic problems with hearing devices reduces malfunction occurrences in the classroom significantly, therefore increasing the classroom competence of hard-of-hearing students. The ability to cooperate with various specialists is also a key factor. Nussbaum (2018) points out that in teaching hard-of-hearing students, one must be prepared for scenarios involving planning and collaborating with medical professionals, educational audiologists, and parents. Another aspect of teaching the hard-of-hearing is differentiation of the lessons to reach the demands of the students (Archbold & Mayer, 2012).

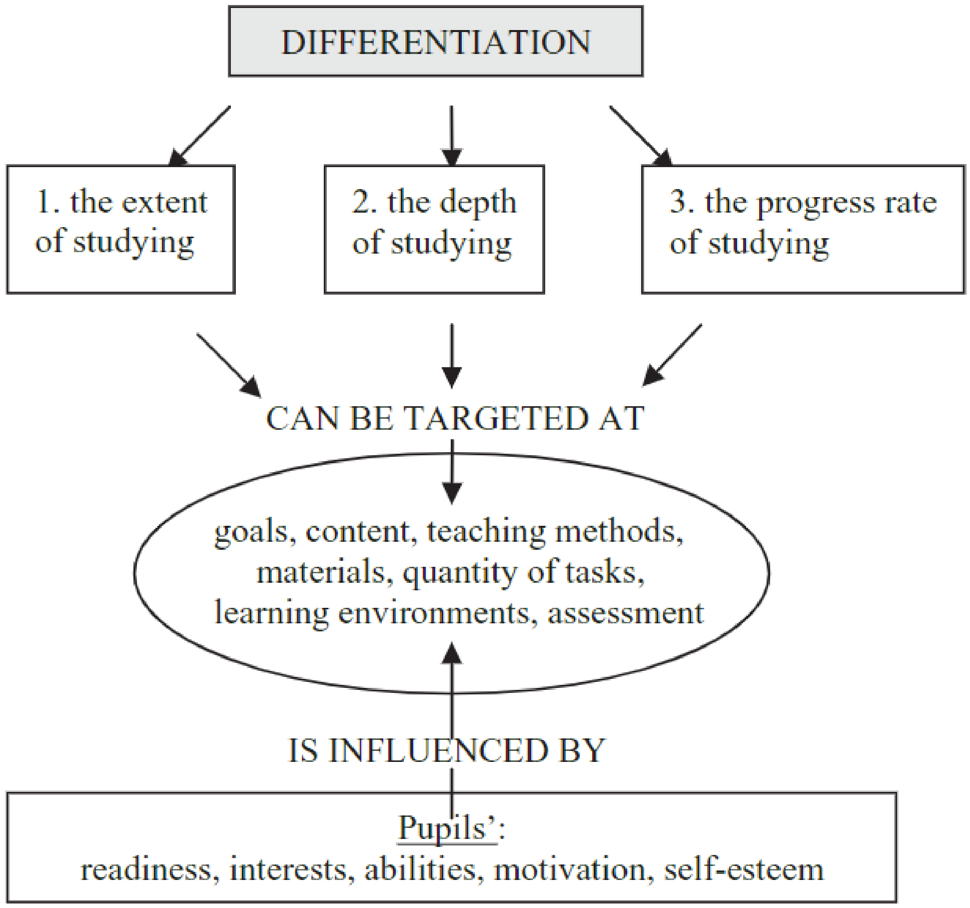

Figure 2: Differentiation can be achieved by altering the extent, depth, and progress rate of studying in accordance with the students' varying demands (Roiha, 2014)

Differentiation for hard-of-hearing students allows to account for the differences in their learning to provide more effective instruction. The principles of differentiation emphasize that the students show variance in their learning preferences, and comprehension (Roiha, 2014). Thus, instruction is tailored to fit the demands of different learning patterns. As demonstrated in Figure 2 (Roiha, 2014), the aspects of a lesson can be changed within the scopes of the extent, depth and progress rate to accommodate the varying properties of a student's learning (UNESCO, 2004).

Archbold & Mayer (2012) specify that contrary to the expectation, the advent of CIs and hearing aids have added a deal of variability to deaf education as the different properties of devices and the users' compatibilities with them can produce a spectrum of learning preferences for hard-of-hearing students. Nussbaum (2018) similarly points out that the differences among the students, their devices and the environments they live in sum up to a wide range of individual difference. Deunk et al. (2018) demonstrate that differentiation practices are most effective when supplemented by measures that consider the environment and individual learners' needs. In this line of reasoning, Nussbaum (2018) prescribes that teachers of hard-of-hearing students carefully consider their preferences and abilities for receiving visual and auditory information, as well as their use of oral and sign language, to inform differentiation strategies.

Another integral component in the mainstream education of hard-of-hearing students is the allocation of resources and planning to create suitable classroom environments. Zanin & Rance's (2016) study finds that students who use hearing complementary devices are at a significant disadvantage compared to their normal hearing peers in the classroom. The usage of remote microphones amplifying the teacher's speech to CIs and hearing aids were found to be effective, and specialized speakers that create "soundfields" which improve the acoustics of the room so that all listeners receive the speaker's audio in an optimized way are recommended. Nussbaum (2018) suggests that the employment of specialists such as sign language interpreters and cued speech translators will improve the accessibility of education to hard-of-hearing students. These improvements can be deployed to improve the quality of education as permitted by an analysis of the benefits and the required costs of implementation.

Some general strategies employed in the mainstream education of hard-of-hearing students include increasing the amount of visual supportive material, using body language extensively, obtaining the students' attention via visual cues before speaking, using open ended questions to check for comprehension, repeating the questions and comments made by students in an adequately audible voice and encouraging frontal seating for hard-of-hearing students (UTSA, 2022; ADCET, 2015). Additionally, advice specific to early language education include increased frequency of role-play activities to encourage dialogue reenactment, games which involve turn-taking and asking and answering of questions, the usage of picture books comprised of pictures from field trips and presentations of such books by telling stories, using art projects and crafts to express language and providing word and vocabulary lists ahead of class time (Manley, 2020; ADCET, 2015). van der Straaten et al.'s (2020) study shows that cochlear implanted children experience a gradual downgrade of quality of life due to lack of social language development, which makes the inclusion of activities which exercise conventions of speech valuable in the classroom.

2.4. Listening Effort

Within this study's scope, listening effort is one of the conceptual cornerstones. In their publication for the Framework for Understanding Effortful Listening (FUEL), Pichora-Fuller et al. (2016) offer the following definition for listening effort: "*The deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task … (which) involve(s) listening.*" Under light of the literature on the nature of listening effort, two thematic perceptions of it arise; the deliberate cognitive effort made to accomplish a listening task, and the sense of exertion and difficulty felt by the listener. The concept of "listening effort" is defined as an umbrella term that encompasses both of these meanings (Lemke & Besser, 2016), which is discussed further in section 5.1.3. *Cognitive Load and Listening Effort*. For the purposes of establishing a general understanding of the concept, listening effort can be explained as the mental effort exerted when completing a listening task (McGarrigle et al., 2014).

It is important to make the distinction that listening effort is a different construct than measures of listening task performance such as speech intelligibility scores and grades of listening test performance (Winn & Teece, 2021). In the study of Pals et al. (2013), it was found that participants who receive improvements to their hearing devices show little to no improvement in their intelligibility measures, while their listening effort measures show decreased mental exertion. By studies conducted on the subject, it is known that people with hearing impairment experience higher levels of listening effort than those with normal hearing (Shields et al., 2022). An increased level of listening effort is connected with heightened experiences of fatigue and stress, and a detrimental effect on the person's effectiveness in terms of multi-tasking and cognitive performance (McGarrigle et al., 2014).

There has been an emergent academic interest in listening effort studies in the past decade, and a variety of ways for measurement have been subject to testing (McGarrigle et al., 2014; Shields et al., 2022). Some of these measures include the testing of cortisol and chromogranin A, self-report measures such as questionnaires and point scales, Functional MRI and EEG readings, skin conductance tests. (McGarrigle et al., 2014; Shields et al., 2022; Pals et al., 2013). Among these, practical and reportedly adequately sensitive measurements are multi-task paradigms with response time measures and pupillometry.

Pupillometry refers to the measuring of parameters related to the eye pupil in response to stimuli. In cases of cognitive exertion, the pupil produces an increase in diameter known as dilation, which is then considered an indicator of mental effort (Granholm & Steinhauer, 2004; Mathôt, 2018). The dilation response of the pupil to a cognitive task is known as a "task evoked pupil response" (TEPR) and has been used as a listening effort indicator (McGarrigle et al., 2020). Zekveld et al. (2010) attest that pupillometry readings of pupil response have been shown to increase with decreased scores of speech intelligibility and increased scores of listening effort. Specialized glasses and visors which record the user's eyes during a task and measure the pupil diameter are used to record TEPR readings and process them as data. As such, pupillometry was chosen to be a measure employed in this study alongside a self-reported point scale and a dual-task paradigm with response time measurement.

Another practical indicator of listening effort is response time. The assumption behind multi-task paradigms is that cognitive effort is split between multiple simultaneous tasks during performance, so measuring performance in two or more tasks at the same time and comparing the results will offer insight on the allocation of attentive cognition (Pals et al., 2013; Pals et al., 2015; Nagels et al., 2020). In this study, the listening test is composed of two simultaneous tasks. While the participant listens to the given audio, they are also instructed to respond to a visual cue. The time taken for the response and the task performance are then taken as response time (RT) data.

For the purposes of this study, I was interested in the relations between possible indicators of listening effort and how they interact during the course of a listening task. The listening test was conceived for this exploration, and I included pupillometry, response time and self-reported listening effort indicator. However, the pupillometry and response time measures were shown to exhibit inadequate reliability, so the conclusions of this study are largely drawn from the self-reported effort scale scores and task performance rates. More information can be found in sections *4.2. Listening Effort Study Design* and *4.4.3. Reliability*.

The studies by Saksida et al. (2021) and Soleymani et al. (2016) show that an impairment to the quality of hearing in early stages of life significantly influences the future development of speech and language. On a more positive note, Waltzman et al. (2003) suggest that the experience of hard-of-hearing children in education might positively affect their language development to reach age-appropriate levels. As Everhardt et al. (2023) display that children with CIs suffer decreased performance the most during listening tasks, an exploration of the effects of listening effort to students' early language education in mainstream schools might offer insight into how future considerations can be shaped to benefit hard-of-hearing students.

3. Aims of the Study and Research Questions

3.1 Research Questions

In consideration of the gap in the academic literature on how listening effort might factor in the language classrooms of hard-of-hearing students and the culmination of previous literature on the subject, this study aims to investigate the effects of listening effort in the early language education of hard-of-hearing students. First, the study aims to gather a deeper insight into the procession of a listening task under an adverse listening condition, and to explore how listening effort experienced by hard-of-hearing students might affect a listening task. Then, bringing the insight from this examination, it is aimed to refer to previous literature to learn more about hard-of-hearing students' experiences in mainstream early language education, and explore possible connections between these accounts and effects of listening effort. Thus, the following research questions are chosen for the study:

1. What is the effect of listening effort on the mainstream early language education of hard-of-hearing students?
2. How do self-reported listening effort, task performance and task difficulty relate in a listening task under difficult hearing conditions?

3.2 Prediction of Results

The study is envisioned to explore the effects of listening effort in a listening task in adverse conditions, and thus on the early language education of hard-of-hearing students in a mainstream school context. The expected result for the quantitative listening task study of the project is that as task difficulty increases, both the task error rate and the perceived effort self-reported scale increases. According to Pals et al.'s (2013) results on a similar premise, it is expected that the increase of the perceived effort rate will be more pronounced that the rate of error, or that the error rate might come to rest at a threshold while the perceived effort rate continues increasing. For the qualitative literature review, the study aims to explore some common themes that arise from the studies of hard-of-hearing students in mainstream education. It is expected that the results and resulting conclusions drawn from the qualitative study will help further explain the themes that arise from the literature review, and some parallels can be drawn between the academic literature regarding listening effort and hard-of-hearing education regarding early language education.

4. Methods

4.1 Research Design

At the conceptualization stages of the study, the aims for the research were designated as the exploration of different methods (pupillometry, reaction times, heart rate variability, etc.) of measurement of listening effort and how well they corelate with the self-perceived effort, while also gaining insight into the conduct of hard-of-hearing students in mainstream early language education. While the specifics of the study evolved during the literature research process (see 3.1 *Research Questions*), the long-standing envisioning for the study was to examine the manifestation of listening effort in an effortful listening task and use the gained insight to commentate on the foreign language classes of hard-of-hearing students. As such, it was decided to adopt an approach that could bridge the quantitative results of the listening effort study and previous literature on hard-of-hearing education.

Johnson et al. (2007) define the mixed methods approach as the type of research in which quantitative and qualitative approaches are combined, in order to widen the horizon of the study and to enhance the understanding provided by the combination of data types. Complementing Johnson et al.'s (2007) definition, Creswell & Plano Clark (2018) offer that the employment of a mixed approach can broaden the width of the understanding provided by an academic study. Cohen et al. (p.31-50, 2018)'s approach to mixed methods research entails that the world may present more than numerical or verbal elements to the researcher, at which time a mixed methods approach may lead to share in these multimodal viewpoints and use it to understand the intricacies of the researched phenomenon. Cohen et al. (p.31-50, 2018) go on to inform about the pragmatic philosophical standpoint of mixed method research, that the meaning is constructed alongside the tangible outcomes from the combination of approaches rather than a sense of argumentative loyalty to a theoretical standpoint. In academic research, there may be certain areas of study that rely on qualitative data sourced from written, spoken, or visual descriptions of events or individuals, while others may utilize quantitative data derived from numerical representations. However, as reinforced by the aforementioned works on mixed methods approach, there are specific contexts and studies where employing both types of data can prove advantageous and enriching. Considering that the aim of this study pertains to the collection and analysis of quantitative and qualitative data, adopting a mixed methods approach was a natural outcome.

Once a mixed methods approach is decided, the next step becomes to determine the respective weights and roles of the qualitative and quantitative approaches. On the question of how the two angles of approach to research can be employed to complement each other, Fetters & Freshwater (2015) introduce the 1 + 1 = 3 framework. The main idea of this approach is that the qualitative and quantitative stages of a research will offer different types of data which upon analysis, will result in diverse manifestations of the phenomenon to be captured and examined. This framework is contested and built upon by the 1 + 1 = 1 framework proposed by Onwuegbuzie & Hitchcock (2019), which emphasizes that an integrated research design will prove more beneficial than a mixed methods study in which the two data types are only used to add width to the exploration. In the 1 + 1 =1 framework, it is proposed that are conducted in a way that their design and analysis processes interact with each other, informing and guiding the conduct of the other approach. Instead of only integrating the two approaches during data interpretation, the 1 + 1 = 1 framework suggests an integrative approach into mixed methods research, in which the intertwined processes of quantitative and qualitative research can contextualize the findings of each other and offer a more holistic view of the subject.
This is argued to produce more connected and integrative results.

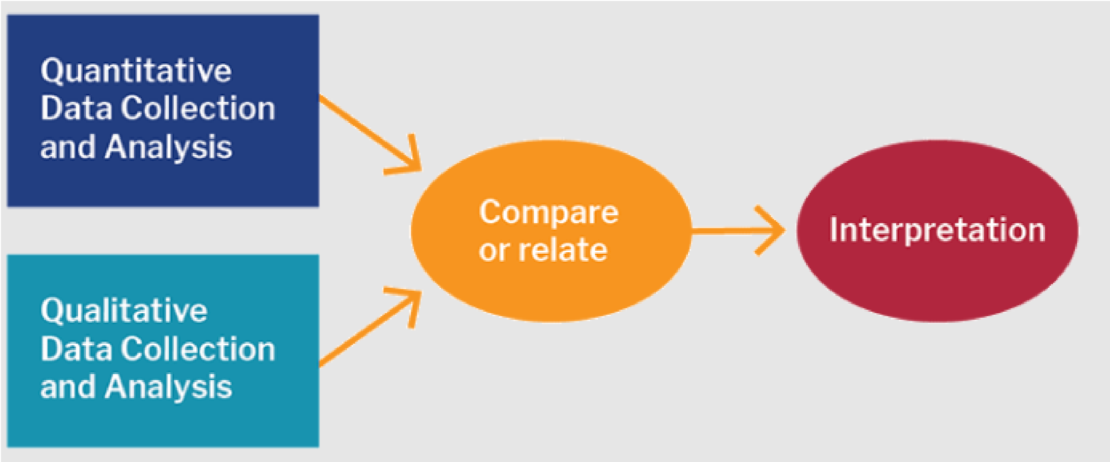

Figure 3: A description of integrated mixed methods approach. The two methods are compared and related to reach an interpretation. (Harvard Catalyst, 2022)

The Figure 3 placed above offers a schema of one such mixed methods study. The study is conducted with interacting quantitative and qualitative stages. The results of the stages inform the planning and analysis processes of each other. The results from both stages are involved in the triangulation of factors contributing to the researched phenomenon, and the integration offers a more in-depth exploration of conclusions. In consideration of the aims and possible methods of this study, I decided to employ the use of the 1 + 1 = 1 framework of integrative mixed methods approach. For the quantitative section, a listening effort study is conducted on participants, data is acquired and analysed to generate insights into how listening effort might manifest in an effortful listening task. On the qualitative step of the study, a literature review is conducted on previous works of hard-of-hearing education in mainstream schooling, and the contextualization of listening effort with the cognitive load theory. The study started with the literature review process, which continued into and affected the conception and conduct of the listening effort test. In turn, the discussion of the literature review was informed and influenced by the results of the test. The conclusions for this study are envisioned to be the results of an integration of the two approaches.

The qualitative part of this study has been the backbone of all considerations since the fundamental beginnings. The research on previous literature and the triangulation of arising themes have been a continuous process throughout all proceedings. The employed method follows a framework of narrative literature review. According Paré & Kitsiou (2017), the narrative literature review is a method which follows an unsystematic rationale of compiling previous literature on a subject. The aim is not to generate generalization, but to summarize the past works and emphasize similarities or common themes. While the criticisms exist that the method tends to be subjective due to its freeform nature, the framework allows for accumulating a synthesis of a volume of literature on a specific subject. The narrative literature review offers pragmatic advantage in providing an overview of information on a selected phenomenon, which suits the aims of this study.

Since the quantitative section of the study follows a more structuralized framework and much of the utilised software tools were custom developed, the design process will be documented in its separate section.

4.2 Listening Effort Study Design

4.2.1 Study Designs in Previous Literature

The listening effort study is essentially a test. The approach I settled on for the design of the listening effort study was first to determine what I was going to test and explore for the purposes of this thesis study (Cohen et al., p.563-564, 2018), and then to observe previous literature and studies to see what kind of approach was taken to accomplish similar purposes. The listening effort study is exploratory in nature and aims to explore the relations and connections between the tested variables of the listening effort perceived by the participant, the participant’s individualized intelligibility levels for easy and hard listening conditions, and the participant’s performance of hearing and repeating the audio stimulus correctly. Due to limited access to participants and equipment, the size of the test group was limited to 5. Compared to the tests in the literature of listening effort, the participant group size of this study is quite small, and while the acquired results might offer insight into the relations of factors in a listening task, should not be generalized into broader populations. The previous literature on listening effort points to pupillometry and response time dual task paradigms as possible objective markers for listening effort measurement (Giuliani et al., 2021; Shields et al., 2022; McGarrigle et al., 2014). In the application of the listening effort study, task-evoked pupil response (TEPR) and response time scores (RT) from a dual task paradigm were also incorporated. The main sources of data of this study have been self-reported effort scores, task performance and grouping by easy and hard difficulties. I have decided not to include the insights from the TEPR analysis as part of the conclusions in this study and have been cautious in reporting the outcomes of the RT scores (for an extensive discussion, see 4.4.3. *Reliability*). Although the collected TEPR and RT data is not considered definitive due to reliability concerns, these measurements form a central part of the listening effort study design and thus will be described in this section.

The main inspiration during the design of the listening effort study has been McGarrigle et al.'s (2020) examination of relations between measurements of pupillometry and self-reported scales of listening effort. The premise of the study was that through pupillometry measurement, the response of the pupil to a listening task (TEPR – task-evoked pupillary response) can be captured and the investigation of TEPR's relations with other measures of listening effort can generate validity to its denomination as an objective marker of listening effort.

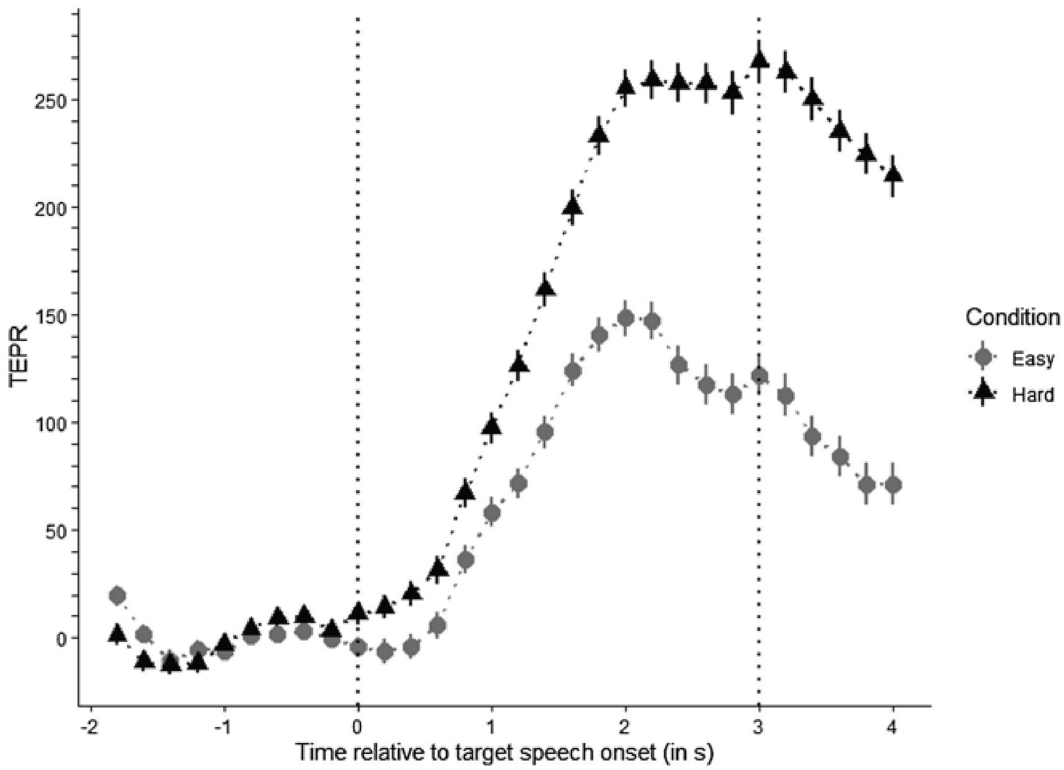

Figure 4: The pupillary response to auditory stimulus during a listening task. The increase in TEPR at audio onset at 0 seconds can be observed (McGarrigle et al, 2020).

In the study (McGarrigle et al., 2020), 28 participants were presented with a listening task and their pupil reactions as well as self-reported perceived effort scores were recorded. The main purpose of the participants was to listen to sentences being read through headphones that were embedded in noise, thus were difficult to hear and understand. The aim of the study was to generate insight into the differences during a listening task in easy conditions versus one in difficult conditions. For this purpose, prior to the main listening task, the participants were held in an adaptive screening listening test. The adaptive worked on the principle of signal-to-noise ratio (SNR), which is a denominator of the difficulty of a listening task. SNR refers to the difference of the heard dB level of the audio stimulus versus the noise it is embedded in. An SNR of -2 dB means that the audio is 2 decibels lower than the embedded noise. During the adaptive test, the SNR for %50 hearing legibility for each individual participant was recorded as the "hard" hearing condition. The "easy" hearing condition was determined as the SNR

of the hard condition, plus 10 dB. For example, a participant who has a hard SNR of -8 dB would have an easy SNR of +2 dB. After determining the SNR for easy and hard conditions, the participants took the main test. They were presented audio stimulus in consecutive easy and hard settings. The results of the study show a higher rate of self-reported listening effort in hard settings compared to easy settings, as well as lower correct response rates and distinctive patterns of TEPR captured by pupillometry measures. The study also found that a more frequent rate of prompting the participants to rate difficulty yielded more sensitivity in self-reported listening effort scale.
In addition to McGarrigle et al. (2020), Koelewijn et al. (2012) and Zekveld et al. (2010) were also guiding studies towards the planning of the listening effort study, as these works also used pupillometry and self-reported effort scales to explore their relations in hearing conditions of varying difficulty. Also being tested on groups of near 30 participants and following a similar study design, their main difference from McGarrigle et al.'s (2020) work is that they Koelewijn et al. (2012) used %50 and %87 hearing legibility thresholds as easy and hard settings respectively, and Zekveld et al. (2010) opted to investigate under three difficulty levels at %50, %71 and %84 hearing legibility thresholds (referred to as SRT in these studies – speech reception threshold).

Alongside self-reported effort scales and pupillometry measures, the listening effort study also includes a secondary visual response task. The idea of using a secondary response task and measuring the deviation from a baseline response time score according to task difficulty can be a method of marking listening effort or the load on working memory, as referenced by several studies (Giuliani et al., 2021; Shields et al., 2022; McGarrigle et al., 2014; de Jong, 2009). This is called a dual task paradigm. As guides for the implementation for this measurement, the focused studies were Pals et al. (2013) and Pals et al. (2015). Pals et al. (2013) conducted the study on 19 participants. The factors measured were the self-reported effort scales and two visual response tasks, as well as the listening task performance. The listening task itself was a simulation of spectral hearing which cochlear implant users experience. The difficulty of the listening condition was determined by the channels used for the sound resolution of the spectral hearing simulation, with more channels being used easing the hearing. The study follows a basis of cognitive load theory as response time (RT) is nominated as a listening effort marker, as resources of the working memory compete for the listening task and response task, an easier listening condition would lead to reduced RTs. The visual tasks employed in the test were a rhyme task and a rotation task. During the rhyme task, the screen would display two word endings in the participants' native Dutch and if the words rhymed, the response button is pressed. For the visual rotation task, the screen would display two Japanese characters with one of them being 90 degrees rotated. If the two characters are the same, the response button is pressed. These visual tasks were presented at the same time as the listening task and the time elapsed between the presentation of the visual stimulus and pressing of the response button was the response time (RT) score. The results of the study showed that as the conditions of hearing were complemented to be easier, the task performance rate increased to reach a certain threshold, while the markers of listening effort showed continued improvement of the participant's experience. In other words, as the conditions were amended, the experienced listening effort continued decreasing even while the correct response rate no longer increased after a threshold (see 6.1. *Listening Effort Study*). Pals et al. (2015) compared the sensitivity of visual and auditory response tasks as part of a dual task paradigm as measures of listening effort. While the study finds that an auditory response task outperforms the visual task, the descriptions of the presentation and the application of the rhyming task (similar to Pals et al., 2013) were taken into account for the secondary response task of the listening effort study.

4.2.2 Study Design

The listening effort study is designed as a dual task paradigm with a primary listening task and a secondary visual response task. For the components of the listening task, the custom Digikuulo software developed by Tuomas Heikka is used. For the response time task and to collect the self-reported effort rating, I've written a complementary program using the Python programming language. During the task, a sample of Tobii Pro Glasses 2 from the Joensuu Campus of University of Eastern Finland Cognition Laboratory of the Department of Computer Science were used to record the pupillary reactions of the participants.

The main listening task of the study is that the participants listen to an audio recording of three digits being read in Finnish. This audio recording is embedded in noise, which is played at 65dB, and the SNR (signal-to-noise ratio) can be manipulated by the program by increasing and decreasing the audio stimulus sound level. This allows for increasing or decreasing the difficulty of hearing. For example, the “hard” listening setting for most of the participants is near -4 SNR, meaning that the audio sound level is 4 dB lower than the noise sound level. The participants’ instructions are to listen to the three digits being played and to repeat them verbally. This is called a “repetition”. In one administration of the test, this process repeats 20 times with a different set of three digits being read each time. For each stage of the study, the participants receive the same digit sets as other participants. In total, one administration of the test consists of 20 repetitions.

Before the test procedure begins, the participants are seated in front of a computer, then they equip the Tobii 2 glasses and a set of headphones. They interact with the program using a standard computer mouse. The researcher is similarly seated in front of a different computer in the same room, although they do not equip any material.

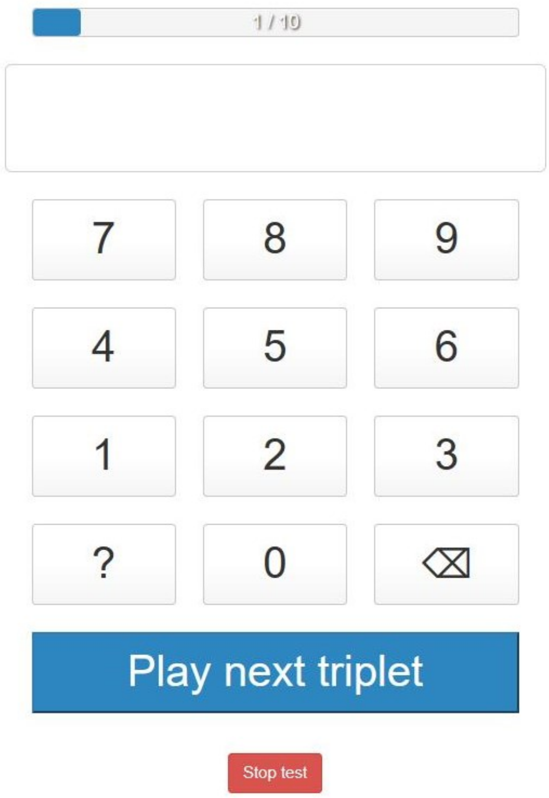

Figure 5: The researcher screen during the listening effort study. The researcher enters the participant's repeated three digits and clicks "play next triplet" to continue to the next repetition.

During the study, the participant and the researcher administrating the test see different screens on different computers. The researcher operates with the screen visible on Figure 5. Once the test is started, the participant hears the audio of three digits and repeats them, which is then entered by the researcher on this screen. This proceed repeats 20 times for one administration of the test.

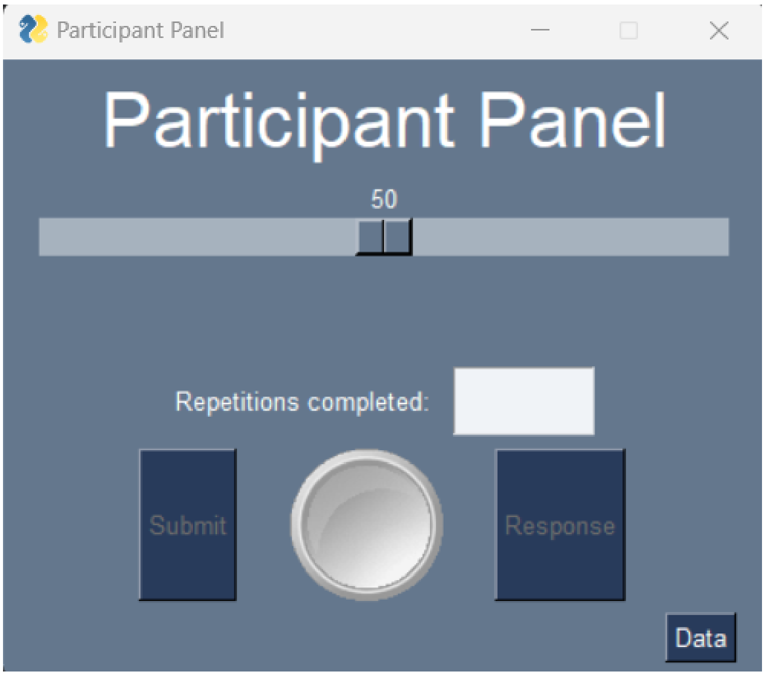

Figure 6: The participant screen. While listening to the digits, the user monitors the button in the centre and clicks "response" when it flashes green. The slider is used to record the perceived task difficulty.

The participant is presented with the screen visualised in Figure 6. Running on a different computer than the researcher's, this program is complementary to the Digikuulo software and uses LAN connection to coordinate the proceeding of the test. This screen is used for the dual task elements of the study, as it allows for the presentation and recording of a response task. When a repetition is started and the audio is played through the headphones, a counter for the response time task also starts. The exact length of this counter is randomized to be between 3.4 and 2.8 seconds, which corresponds to end during the playing of the audio recording. When the counter ends, the grey button in the centre of the screen turns green. The participant is instructed to simultaneously monitor the button in the centre while simultaneously listening to the audio. Keeping the mouse cursor over the "Response" button on the right, once the button turns green, the participant clicks. This records the time elapsed between the appearance of the green light and the pressing of the response button. The screen also features a self-reported effort scale which is administered as a slider -

from which is recorded and corresponds to the subjective rating of listening difficulty -, and a “Submit” button.

During a repetition, the participant performs four actions. First, while the audio of the digits is playing, they monitor the button and click “Response” when it turns green. Then, they repeat the digits they’ve heard. Third, they use the slider on the upper half of the screen to give the repetition a rating of how difficult it was for them, from a scale of 0 to 100. Finally, the participant clicks the “Submit” button to proceed to the next repetition.

In a session of the listening effort study, the test is administered in differing settings for a total of four times. The order of these four steps is as follows: the practice test, the response time baseline test, the adaptive test and the alternating difficulty test. In the practice test, the participant completes 30 repetitions. The SNR is 0, so the noise and audio stimulus are presented at the same sound level. The participants are introduced to the four actions they will perform during testing one by one. For the first 10 repetitions, they only listen to the digits and repeat what they’ve heard, then press the “Submit” button. For the next 10 digits, the response task is introduced. They click “Response” once they see the green light, repeat the digits, and press “Submit”. For the last 10 repetitions, they also use the slider to record their subjective perception of listening difficulty. They click “Response” once they see the green light, repeat the digits, rate the difficulty experienced during repetition, and click “Submit”. The data from this step is not recorded and it is administered as a way to introduce the study design and the dual task paradigm to the participant. By the end of the practice test, the participants were accommodated to the simultaneous nature of the tasks and the order of their actions.

The second step is the recording of the baseline for response time scores. The participants wear no equipment, the listening task is omitted, and the only action is to monitor the button and click “Response” once it turns green for 20 repetitions. The participants are instructed to click the response button as quickly as possible once the visual stimulus is presented. The average of the response times (RT) from this administration is later used as the RT baseline.

The third step is the adaptive test. The purpose of this section of the study is to determine the individualized signal-to-noise ratios to create easy and hard difficulty settings for each participant. As the participants' hearing levels and the levels of familiarity with Finnish might affect the results, this process is integral to determine the same difficulty classification for individual participants. The participant wears the equipment and completes 20 repetitions, performing all four actions and all tasks. The program records the participant's repetition of the digits and if there is an error, the SNR is increased, meaning that the audio stimulus dB is increased by 2 to be better heard over the noise. If the participant's answer is correct, the SNR is decreased by 2, so that the next repetition becomes more difficult. Over the course of 20 repetitions, the program determines the corresponding SNRs to the hearing legibility ranges of 50% and %90 (referred to as L50 and L90) based on the calculation of the participant's answers (for more information on the calculations employed by the Digikuulo software to determine hearing legibility ranges, see Willberg, 2016). These SNR settings are used in the next step to determine the easy and hard listening conditions.

The final step is the alternating difficulty test. This is the main test of the session, in which the data is recorded and collected for analysis. The participants wear the equipment and complete 20 repetitions with all tasks included. In this test, the digits are played in pairs of two easy and two hard repetitions. In total, the participants hear ten easy repetitions and ten hard, alternating in pairs. The easy setting is the SNR corresponding to L50, and hard is the SNR corresponding to L90, as determined by the adaptive screening in the previous step. In the study, the easy setting simulates a normal hearing experience in the listening task, while the hard setting represents the experience of listening in noticably adverse conditions, such as that experienced by cochlear implant users. The main data collected from this listening effort study is the self-reported effort scores and the rate of correctly repeated digits. This data is compared based on easy and hard listening conditions, and the displayed differences and relations are anticipated to be representative of the same listening task completed in easy and hard conditions, simulating the difference between normal hearing, and hearing experienced with cochlear implants.

The study procedure was first prototyped during the preliminary testing phase in August to June 2023. Afterwards, the procedure was finalized as the four step administrations described above, with improvements being made to the program to better coordinate with the Digikuulo, for practicality during test administration and better data processing. The procedure consists of greeting the participant, presenting the ethical consent form, providing information about the aims of the study and instructions of the testing tasks, the administration of the four steps of the study and debriefing. The test typically takes between 40 minutes and 1 hour to administer and complete.

4.3 Data Collection

4.3.1 Literature Review

The data for the literature review was collected from PubMed, a comprehensive database of biomedical literature, as well as UEF Primo, which allowed access to a range of educational and methodological resources. The sources gathered were categorized under three sections which also correspond to the sections of the literature review itself: historical framework of hard-of-hearing education, language development of hard-of-hearing children, and cognitive load theory and contextualization of listening effort. There were 41 resources analysed and discussed in total. The distribution of the resources for sections are as follows: 12 for historical framework of hard-of-hearing education, 12 for language development of hard-of-hearing children, and 17 for cognitive load theory and listening effort. The collected data was synthesized and organized in a narrative structure.

4.3.2 Listening Effort Study

The listening effort study was conducted on two testing phases. The first phase took place during August-June 2023 and was conducted for the purpose of acting as a preliminary stage, in which the test procedure was tested, and the shortcomings were noted to be filled. Six participants were enlisted from the Joensuu area. The participants were either university students or research personnel, between the ages of 20 to 31. Prior to the test, the following timeline was created, and the test process was continued accordingly:

1. The participant enters the room and is greeted by the researcher.
2. A copy of the consent form is produced and given to the participant.
3. The main points of the consent form are explained verbally, and time will be given for the participant (about 5 minutes) to read the contents.
4. The participant is asked for any questions they might have, and they are answered.
5. The participant declares consent to participate in the study.
6. The background interview begins. The participants are asked about their age, gender, occupation, their major at the university if students, hearing-related medical background, their native languages, and their level of Finnish (little to no aptitude – beginner – intermediary – fluent).
7. The software and equipment are prepared.
8. The participant is informed about the tasks of the study.
9. Calibration for the Tobii Pro Glasses 2 is made.
10. The study test procedure takes place.
11. After the study, a debriefing interview is made with the participant to learn which parts of the design they thought were difficult, and what factors challenged them if any.
12. The researcher expresses gratitude for the participant's presence, and the study is ended.
13. The researcher saves the acquired data.

Through the findings of the preliminary testing and the feedback from the participants, improvements were made to the user interface and complementation of the Digikuulo program. The second phase of testing was administered and completed during December 2023 to January 2024. Five participants were enlisted from the Joensuu area, all university students aged 20 to 28. The same procedure described in the timeline was administered. During February of 2024, two of the five participants from the second testing phase were recalled and administered the test again, to gather insight on the reliability of the test measurements. The duration of the administration of the test to individual participants ranged from 40 minutes to 1 hour. I have conducted all tests in the University of Eastern Finland Cognition Laboratory of the Department of Computer Science. The Tobii 2 glasses, the Tobii 2 software, and the computer for recording the data with the glasses were the courtesy of the Cognition Laboratory. The headphones, Digikuulo software and the computer used by the participant were supplied by the Kuopio University Hospital. While the experiences during preliminary tests were used as guides to finalize the prototype of the test procedure and program, only the data from the main testing phased during December 2023 – January 2024 were used in the data analysis.

4.4 Research Considerations

4.4.1 Research Ethics

The guidelines for the ethical considerations in this study were the code of conduct published by the Finnish National Board in Research Integrity (TENK, 2023) and Cohen et al.'s (pp. 111-143, 2018) chapter on the ethics of educational research.
During this research, it was made primary concern to inform the participants towards the purposes and conduct of this research to the best of their knowledge and to answer any questions they might have beforehand to acquire their informed consent, to take the necessary measures during recording and reporting of data to ensure the continued anonymity of the participants, as well as ensure the principles of trusted confidentiality,

to ensure the omission of any information that could lead to identification of personal details of the participant which could lead to a breach of non-traceability, to ensure that the procedure and reporting of the collected data lead to no detriment to the direct or indirect participants, and to protect the utmost measures of privacy (Cohen et al., 2018).

To ensure that the data collection adheres to the safeguard principle described in TENK (2023), all participants have been informed of the purpose and progression of the study and have received letters of ethical consent, which they've accepted of their free will. The ethical consent form was authored in consideration of the UEF guidance on research ethics (UEF, 2023), and informed consent was elicited by disclosure of the purposes and methods of research. Before declaring consent to participate in the study, the participants were read aloud a summary of the ethical consent form, as well as being allocated time to read the document on their own. The participants were asked if they had any questions about the nature of the study and it was made sure that they had no further inquiries. The participants were sought to be treated respectfully and appreciative of their participation. To adhere to the principle of plagiarism (TENK, 2023) it was taken into care that all previous literature including text, ideas, results, figures, and tables were properly referenced and attributed to their original source.

4.4.2 Validity

Validity refers to the extent offered by the academic research to which the insights generated from the study can be attributed to the understanding of the phenomenon to which they aim to explore (Cohen et al., 2018). This section will be concerned with the investigations of the internal, external, construct, and content validity in regard to this study. The following descriptions of the types of validity and the guidelines for the measures employed in this study to mitigate invalidation are derived from the chapter on research validity by Cohen et al. (pp. 252-267, 2018)

Internal validity is the degree to which the measures employed in the research securely relate to the phenomenon in question. A case of high internal validity can be achieved by ensuring that the factors which contribute to the measured attributes are accounted for in the research. External validity refers to the degree to which the findings and insights of the study can be generalized to a larger population or varying settings. The external validity can be sought by minding the range of generalization considering the size of the sample examined and the population that can be generalized to. Construct validity is the quintessential form of validity that is sought in research, as the degree to which the methods used in the study are related to the examination of the selected constructs. Concept validity is the range of the phenomenon that the testing instruments can cover.

To ensure internal validity, the listening effort study was conducted in a closed space, with no visual or auditory distractions for the participants. The only stimulus for the participants were provided through either the computer screen with a black background, or soundproof headphones. For the literature review, the sources were sought to be selected by a comprehensive search, included in regard to their relevance and cumulative contribution to the subject of the study.

Due to the limitations of the study, the external validity of this study must be supplemented in accordance with previous literature. Since during both data collection methods the sample size was quite small compared to the usual sizes of participant groups for similar types of studies, generalizations must be made with caution. During the conclusion section, the insights gathered from the studies were only taken into account if they had supported previous findings on similar studies with larger sample sizes.

The construct validity was ensured by the phase of literature review and theoretical research conducted throughout the data collection of the study. The decisions on determining the methods and instruments of collecting data were made by referring to either the background of literature or to methods used in previous studies during the phases of research design and data collection.

Concept validity in the case of this research can be interpreted as the extent to which the differences between easy and hard hearing conditions can be displayed in twenty repetitions for one admission of the test, and the extent to which the literature review covered the phenomenon of hard-of-hearing children's experiences in early language education. The number of twenty repetitions was chosen as it was deemed to be extended enough for valid generation of results and short enough that the participants would not experience extended amounts of fatigue and it would not affect the internal validity. For the literature review, three main topics of interest were designated, and 41 guiding sources were selected to both provide depth and also cover a wide selection of aspects regarding the study subject.

In summary, the validity of this study has been considered and effort was addressed through various measures aimed at ensuring the legitimacy of the findings. These efforts were exerted for the purpose of enhancing the overall validity of the study, providing a robust foundation for interpreting the insights generated and their contribution to understanding the effects of listening effort on the experiences of hard-of-hearing children in early language education.

4.4.3 Reliability

Reliability refers to the exactness and precision of the instruments and methods used in an academic study (Cohen et al., pp.268, 2018). In the literature review section of the study, consideration was taken that the subjectivity of qualitative analysis was contained to the boundaries of the study and relevance to the study topic was sought. For the listening effort study, two of the five participants (participants 9 and 11) from the December 2023 – January 2024 testing phase were re-invited to take the test after one month, so that the test-retest reliability of the study can be measured (Jhangiani & Price, 2017).

To measure reliability, the datasets of the same participants' first and second test administrations were compared in a correlation matrix. A higher Pearson's r value and therefore a lower p value would indicate a correlation between the two administrations, showing that the test produces similar and stable results under repeated circumstances.

Table 1: Correlation matrices of RT and TEPR scores, among two administrations of the test.

Participant	Measure	Pearson's r	p-value
Participant 9	RT	-0.073	0.761
Participant 11	RT	-0,048	0.841
Participant 9	TEPR	-0.064	0.790
Participant 11	TEPR	-0.181	0.446

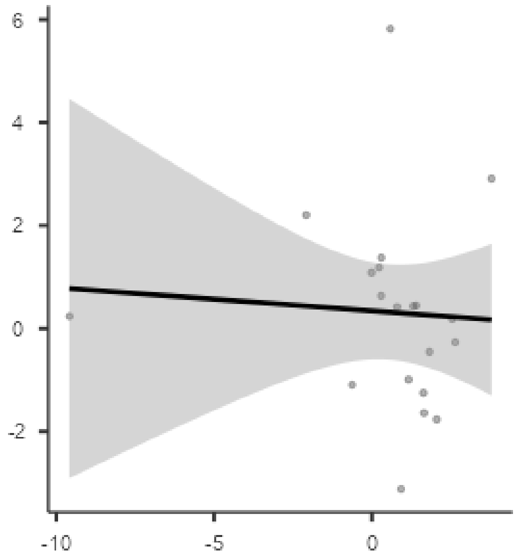

Figure 7: Correlation figure from Participant 9's dataset. Horizontal: TEPR score from session 1. Vertical: TEPR score from session 2.

According to the correlation tests, the TEPR and RT scores show little similarity among consequent administrations. For both participants, the comparisons between the two datasets of TEPR and RT scores display high p values, indicating no correlation between the two test admissions.

The same test was applied to the self-reported perceived effort and digit performance measures. Pearson's r was calculated for self-reported perceived effort scores as the data is normally distributed, but for digit performance Kendall's tau-b was calculated as the scores of 0-4 are not normally distributed and ties in ranks can be present.

Table 2: Correlation matrices of self-reported effort ratings and digit performances among two administrations of the test.

Participant	**Measure**	**Pearson's r / Kendall's tau-b**	**p-value**
Participant 9	Self-reported perceived effort	0.396	0.084
Participant 11	Self-reported perceived effort	0.298	0.201
Participant 9	Digit Performance	0.344	0.112
Participant 11	Digit Performance	0.000	1.000

Note. Pearson's r used for self-reported perceived effort, and Kendall's tau-b for digit performance.

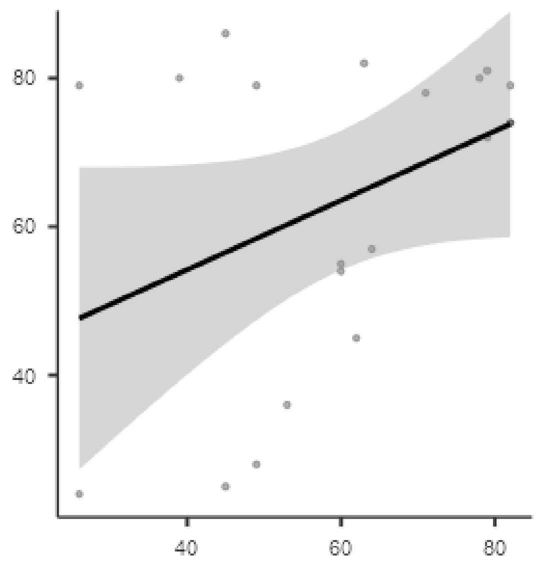

Figure 8: Correlation figure from Participant 9's dataset. Horizontal: self-reported effort rating from session 1. Vertical: self-reported effort rating from session 2

The self-reported effort rating and digit performance comparisons among administrations display desirable results, with lower p values and positive correlations, with the exception of Participant 11's digit performance correlation. The dataset is comprised of only two participants and a wider participant pool would be more robust to gather insights, yet the data shown suggests that the study did not reliably measure TEPR and RT scores. Participant 9 confided that they had concentrated very much on the response time task on the first session of the study, and on the second time they deliberately focused more on the listening task to be more accurate. The discrepancy could be caused by this effect of deliberate allocation of working memory to the secondary task rather than to the listening task or could be due to possible errors of the dated pupillometry device. Due to the insufficient reliability, TEPR and RT data were not added to consideration during the discussion section, although the steps followed in obtaining them were reported in the data analysis sections. The reliability matrix test showed desirable results in the self-reported effort scale, so the data from the effort slider and the task performance were included as reliably acquired data in the conclusions of the study.

5. Analysis

5.1 Literature Review

The study of the measurement and effects of listening effort is a relatively recent trend in research, and thus its proximity to language education has not yet been supported by prominent literature. Since this study is focused on the exploration of the effects of listening effort in the early language education of hard-of-hearing children, the foundational theoretical framework derived from the literature review is split under three headlines: the historical framework for providing education to hard-of-hearing students, the language development of hard-of-hearing children and the cognitive load theory and how listening effort can act as an indicator of mental exertion.

5.1.1 Historical Framework of Hard-of-Hearing Education

Lang's (2011) chapter in *The Oxford Handbook of Deaf Studies, Language, and Education* paints a picture of two conflicting views throughout the early history of deaf and hard-of-hearing (DHH) education. On one side is the proponents of manualism, the method of instruction in which sign language is used as the primary input in the schools for hard-of-hearing students, while also being acknowledged as an element of culture. On the other, the advocates of oralism, who argue that an early exposure to vocal language and development of ways to integrate into it – such as lipreading – must be the cornerstone of DHH education. The debate eventually reached its culmination at the historical 1880 International Congress of Education in Milan, where the elements of manualism in DHH education were agreed to be limited, and oralism was set to be the dominant educational approach in Europe (Lang, 2011).

While the manualist practice of early exposure to sign language is shown by modern research (Goldin-Meadow & Mayberry, 2001; Lang, 2011) to predict future academic success and literacy level, the resolutions of the Milan congress took hold and negatively impacted DHH education, leading to bans of sign language in deaf education institutes and deaf educators struggling to find a place in the education of their own

communities (Lang, 2011). The notions of the Milan conference were rejected in the 2010 ICED conference in Vancouver and the recognition of sign languages as native languages of the DHH students were encouraged (Moores, 2010; ICED, 2018).
While the final official amendments were made more than a hundred years later, alternative approaches that took into consideration the value and effect of sign language challenged the oralist approach starting much earlier. During the 1980s, an approach advocated for by Marie Jean Philip was the bilingual-bicultural education (Bi-Bi) (Bienvenu & Meehan, 2022). This approach emphasized the importance of sign language as a native language to the deaf and hard-of-hearing children, advocating for its placement as a primary mode of instruction in their education. The approach also made a push to view deafness as a cultural identity and to include the students as part of this community (Bienvenu & Meehan, 2022). Another alternative method was total communication, which held the idea that visual, oral and manual modes of communication must be integrated into DHH education in order to provide the optimal social and educational benefits to students (Lang, 2011). The National Association of the Deaf in the United States supports a dual language development approach that draws from the ideas of the total communication approach (NAD, 2020). It is clear from the examination of these sources that in DHH education, the opportunity to learn and communicate through sign language carries utmost importance; although an overview of DHH education would not be complete without accounting for the impact of technological advancements and mainstreaming that developed through the last twenty years.

The advancement and more common use of technological means such as hearing aids and cochlear implants has led to a pivotal shift in the education of deaf and hard-of-hearing students. The reports of recent studies support that especially with early implantation, children with moderate to severe hearing loss can continue their speech development alongside their peers (Sharma et al., 2020; Lonka, 2014, p.19-23; Geers et al., 2003), which leads to many of them being included in mainstream education scenarios (Pritchard, 2013, p.116). Geers et al. (2017) find that the best results for hard-of-hearing children are reached with no input of sign language and cochlear implantation at the earliest opportunity. The recommended course of educational planning for hard-of-hearing children is set to be early intervention and administration of auditory-verbal therapy alongside exposure to any form of language in early years and ultimately inclusion in mainstream education with the consideration of the child's potentially varying needs (WHO, 2018; U.S. Department of Health and Human Services, 2024; Allen & Morere, 2020; Nadege et al., 2011).

While mainstreaming for hard-of-hearing students has the benefits of supporting successful development and further predicting success in continuous education (Nadege, 2011), it is not without challenges. Mathews (2017) argues that mainstreaming can have negative effects on the hard-of-hearing students by creating situations of social stigma and isolation. As FEADPA (2023) attests that hearing impairment experienced in childhood impacts social development, this is an important consideration in the education of hard-of-hearing children. It is also a known fact that hard-of-hearing students are a group with high degrees of variance, stemming from the devices that they use, their environments and the students themselves (FEADPA, 2023; Archbold & Mayer, 2012; Nussbaum, 2018). An example to how these challenges might be navigated is within the Finnish education system, in which the CI students from this study are also accommodated. The three-tiered support system forms the backbone for special needs support in Finnish schools. This system allows the school decision making authorities as well as the teachers and supporting staff to track the students' individual levels of need for support and work together to provide the required services and accommodations (Sundqvist et al., 2019). A display of this system can be seen in Lonka's 2014 dissertation, in a section where the population of the pupils that were designated as participants for the study is described. The hard-of-hearing children that were part of mainstream education are reported to be receiving support from classroom assistants, sign language interpreters, special teachers or benefiting from individualized plans, depending on their individual needs (Lonka, 2014, p. 43).

In tracing a short summary of DHH education it becomes evident that the field has undergone significant transformation over time. While the era before the advent of technological hearing devices is characterized by the struggle to establish sign language as an element of instruction, language, and culture; the modern discourse centers around the importance of early intervention, adequate and individualized support for the students and traversing the course of mainstreaming with its challenges and benefits. By heeding the findings on the continuation of DHH education, educators and policymakers can continue to advance the quality and effectiveness of education for hard-of-hearing students.

5.1.2 Language Development of Hard-of-Hearing Children

One of the conclusions of the previous section on the historical framework of DHH education is that the hard-of-hearing students form a population far from homogeneity. The differences in the individual child's level of age-appropriate language development are informing factors in the decision-making processes regarding their education and the means and level of support needed for their wellbeing. In order to better understand the dynamics of hard-of-hearing students in mainstream foreign language classrooms, it is necessary to examine the factors which impact early language development.

According to the National Institute of Deafness and Other Communication Disorders' pamphlet "Speech and language developmental milestones" (NIDCD, 2010), the first three years of life constitutes a prime period for the acquisition of the speech and language abilities required for mastery of the first language. The typical language development for children in this period is characterized at first by reactive behaviours such as turning to direction of sounds, reacting to speech of parents and responding to changes in voice tone (ASHA, 2023). With growth during and after the first two years of age, children are observed to babble, point to parts of the body or object when asked, follow simple instructions, and utter increasingly longer strings of words (ASHA, 2023). By three years old, normally developing children achieve understandable speech (ASHA, 2023). Hearing impairment experienced in these early stages of development can cause any selection of these milestones to be delayed, and future linguistic development to be hampered.

For the cases of children who show delayed linguistic signs, several methods are used to determine the level of language development. The methods of testing focus on some markets of language development and use an inventory of items to test the level of performance which relates to these markers. Aguilar-Mediavilla et al. (2022) list some of these markers to be phonetics, vocabulary knowledge, syntax competency and body gestures. The MacArthur-Bates Communicative Development Inventories (MBCDI) are a set of such tests which measure the level of language development of a child (Marchman & Dale, 2023). The test is taken by the parent and their reports are measured up to the normal development standards of development markers to determine the level of the child's linguistic capabilities (Marchman & Dale, 2023). While the MBCDI are a commonly used test in clinical research settings and are adapted to several languages, it is not rare to observe similarly developed alternative tests in various studies.

Another method of measuring language development is through tests which focus on phonological awareness. It is assumed that phonological and lexical competence of children are indicators of their level of language development, and studies show that ties are established between the phonological awareness and phonological working memory (Rodrigues & Befi-Lopes, 2009). Under these findings, nonword tests can be encountered as measures of phonological working memory and therefore the level of language development (Gathercole et al., 1994). A commonly utilised instance is a nonword repetition test. In these tests, children listen to nonwords which adhere to the appropriate morphological rules of their native languages and are asked to repeat them as accurately as they can (Gathercole et al., 1994). A variant of this for measuring reading development is the nonword reading test, in which the children are asked to read out loud pieces of text made from nonwords (Colenbrander et al., 2011).

Children who experience hearing loss during early stages of development show lower measurements of the markers of language development mentioned above. Briscoe et al. (2001) find that children with mild to moderate sensorineural hearing loss (hearing impairment stemming from the inner ear) experience significant problems in nonword repetition and phonological competence tasks. American Speech Language Hearing Association's pamphlet "Effects of hearing loss on development" (ASHA, 2015) informs that early hearing impairment causes delayed development of speech and communication skills, slower development of vocabulary, increased difficulty in comprehension of words with multiple meanings and sentences of high complexity of structure. These findings are also supported by the longitudinal study of four CI children (Moeller et al., 2010), in which they were observed to be at risk of delayed language development in several areas compared to their normal hearing peers. It should be noted that the results from Moeller et al. (2010) and Briscoe et al. (2001) are derived from children who were implanted at a late age, typically after 2-3 years of age. In contrast, the measurements of language development of early implanted children paint a more optimistic picture.

Cuda et al.'s (2014) study on the effects of early implantation on pre-school children present interesting findings. It was found that the children with no impairments other than hearing loss develop language skills at an "almost normal" rate when implanted before 12 months of age. Shojaei et al. (2016) conducted a study comparing the language competences of children implanted before 6 months with those implanted between 6 and 12 months, and the results are equally interesting with Cuda et al.'s (2014) study. The early intervened children displayed higher performance than the late implanted group in all assessments. These findings are supported by similar works (Sharma et al., 2020; Lonka, 2014, p.19-23; Geers et al., 2003) which show that early implanted CI children develop language to a level that is close or parallel to normal hearing children. Another intriguing report on the subject comes from an examination not of first language acquisition, but of foreign language learning. Everhardt et al.'s (2023) recent study has compared the foreign language competencies of two groups of upper secondary level students: one being an early cochlear implanted group and the other a group of normal hearing students. The results of the two groups were shown to be statistically paralleled in measurements of reading skills, while showing a discrepancy in listening skills, where the CI group underperformed. These findings show that with early intervention and necessary support, CI users can develop language adequately to their age-appropriate level and use their skills towards learning a foreign language in mainstream education. The area of challenge for the CI group in Everhardt et al.'s (2023) study seems to be listening, which prompts for an investigation of how listening effort might affect the performances of hard-of-hearing students.

5.1.3 Cognitive Load Theory and Listening Effort

With the descriptions offered in the section 2.4. *Listening Effort*, one can summarize listening effort as the mental effort experienced while completing a listening task. While this summarized definition can help conceptualization, an overview into the cognitive load theory is needed for clarification of the principles of listening effort. The literature on cognitive load theory has developed in close relation with educational methodology, as the theory itself was aimed to understand and explain the working principles of human apprehension, and by these insights create guidelines and methods to increase effectiveness of instruction (Hanham et al., 2023). The theory has its origins in John Sweller's work in the 80s (CESE, 2017). Studying the ways of instruction in teaching complex themes, Sweller (1988) found that learners could manage to solve a problem, but they seemed to fail at applying the problem-solving strategy in diverse settings and using it across transformed types of the same problem. He concluded that the over-reliance on problem solving methods induced a high cognitive load on the students. The cognitive load theory has its basis on the foundation that the human learning capacity is composed of a long-term and a short-term memory (de Jong, 2009). The short-term memory is also called "working memory" in order to clarify that it is in charge of information processing. While the capacity of the long-term memory is regarded as limitless, the working memory is bound by a certain limit to the amount of information that can be stored (de Jong, 2009; CESE, 2017). The underlying principle of the theory is as such: in cases where a learning task and the instructional methods of presenting the task create excessive cognitive load, the working memory tends to be exhausted before the information is processed and learning can take place.

In studies regarding cognitive load, there are mainly three categories of approaches to measure the load created on the working memory: questionnaires and scales used by the participants for self-rating; physiological markers such as pupillometry, heart rate measurements, fMRIs and EEG scans; and dual task paradigms in which the load is measured by the variance of performance for the main or secondary task (de Jong, 2009). The most commonly used measurement is the self-report scales, on which Ouwehand et al. (2021) report that point-based scales such as sliders and Likert scales are more suited for complex tasks, and scales with visual indications or pictorial elements are better suited for studies involving simpler tasks. Drawing from the report on measurements of listening effort from section 2.4., it can be noted that listening effort and cognitive load share the same methods for measuring their effects on participants in studies. This is further evidenced by the bridging study by Marsella et al. (2017), where the EEG scanning results of participants during effortful listening indicate increased cognitive load.

The research on cognitive load leads to using the gained insights towards enhancing instruction in educational settings. The instructional methods stemming from cognitive load aim to decrease the load on the working memory and allow for more mental resources to be allocated for learning. As discussed before, one finding of the research on cognitive load is that focusing on teaching problem solving strategies and leaving learners to apply the strategies to solve transformed problems creates high cognitive load. As such, it is recommended to employ "worked examples" where learners are presented with various types of problems solved by expert methods, before they are expected to apply the strategies themselves (Paas, 1992; Hanham, 2023; CESE, 2017). Another finding is that the working memory performs better when refreshed through rest, so short and regular breaks are suggested for a reduced feeling of tiredness and more efficient learning (Hanhan, 2023). A concept that rises from cognitive load theory is modality, which shows that the working memory operates on two subprocessors; one based on visuospatial information such as diagrams and written material, and one of phonological/verbal information, which is involved in comprehending either written or spoken language (Hanham, 2023; de Jong, 2009). Studies conducted on this concept show that when the incoming information is shared by the two subprocessors, the working memory load is more manageable. An example is Mousavi's (1995) study where two groups of students were asked to solve geometry problems. Prior to the task, the first group received written instructions composed of text and diagrams, while the second group received the same diagrams, but the text instructions were replaced by an audio recording. The audiovisual instructed group surpassed the visually instructed group in performance. As a general rule, the cognitive load theory suggests that working memory capacity spent on adverse conditions of retaining information impairs the learning efficiency and task performance. Remembering that studies on listening effort suggest more mental exertion is observed in more adverse hearing conditions (see section 2.4. *Listening Effort*), the two topics of research draw on a similar theme.

The similarities between listening effort and cognitive load are further pronounced by Paas's (1992) classification of two components of cognitive load: mental load and mental effort. He defines mental load as the difficulty imposed by the was information is delivered to the learner, and mental effort as the exertion of the learner's mental capacity to accomplish the task. A parallel between listening effort and cognitive load can be drawn on the account of the similar explanation that mental resources allocated to the reception of information overburdens the capacity and leaves less for the cognition of information (see Hunter & Pisoni, 2018). Lemke & Besser (2016) also pronounce this similarity and go on to construct listening effort as a combination of two components, influenced by cognitive load theory. The first is "perceived effort", the sense of difficulty and discomfort experienced by the listener during a listening task, that is induced by adverse conditions of hearing. The second is "processing load", the allocation of additional cognitive resources for the completion of the listening task. This model of listening effort inspired by cognitive load theory is supported by Hunter's (2021) findings which denote that speech perception in tasks which are embedded in a difficult listening environment absorb working memory capacity which could otherwise contribute to effective multitasking. The discussion and findings on listening effort and cognitive load studies show that hearing impairment is not an isolated medical problem, but also a factor which significantly influences cognitive operations, which are routinely used in educational settings (Peele, 2018).

From its conceptualization, the cognitive load theory has been influencing educational methodology. The main applicable counsels of the studies on cognitive load have been to align the educational content with the pre-existing knowledge of the learner to facilitate the forming of connections, to configure the delivery of information in a way that relieves unnecessary cognitive load placed on the working memory, and to stimulate the activation of knowledge in ways that the learner can develop expertise over following worked examples (de Jong, 2009). In the case of early language education, Handley's (2021) article discusses a similarity between chess and language classes on the basis of cognitive load. Just as a chess player builds expertise by repeating the same patterns to familiarize with different situations in games, a student of language requires acclimatization to concepts and vocabulary before being able to function in communicative tasks. The working memory can be complemented by completing reading and vocabulary tasks before working on complex tasks, as familiarization with the material will relieve the need to decode chunks of language word by word (Handley, 2021). In Lynch's (2015) study, the international students' reports and concerns over difficulty following lessons in English are compiled. Similar to Handley's (2021) ideas, the suggestions for lecturers from Lynch's (2015) study include reducing the quantity of material presented during the lecture and emphasizing pre-lecture and post-lecture material, controlling the speed of speaking, providing visual support, and having a relaxed classroom atmosphere to encourage asking of questions and asking for clarification. While the cognitive load on listening in hard-of-hearing children might be higher due to listening effort, using the principles of cognitive load theory to structure lessons can help mitigate the fatigue and frustration that can build over the increased effort of listening.

5.2 Listening Effort Study

The data presented in this section is gathered from the five participants who were enlisted during second testing phase from December 2023 – January 2024 (more information about the participants can be found in section 4.3.2 Listening Effort Study). The participants followed through the four-step structure of the test procedure, which are ordered as the practice test, the response time baseline test, adaptive test and alternating difficulty test. The presented data is from the final step in which the participants completed the test under easy and hard conditions. More information on the determination of easy and hard task difficulties, the test procedure and timeline of the study can be found in sections 4.2. *Listening Effort Study Design* and 4.3.2 *Listening Effort Study*. To gather insight on the reliability of the test procedure, two of the same participants were invited again during February 2024 and were administered the test once more with the same material and equipment. Navarro & Foxcroft's (2022) resource on statistical tests and their applications via jamovi (The jamovi project, 2024) software was used for all decision-making processes including normality tests, using paired samples tests using statistics which do not assume normality for digit performance scores and methods used for test-retest reliability.

Through the administration of the alternating difficulty test, the participants completed 20 repetitions of a listening task and data is collected for each repetition from four sources: the self-reported degree of experienced difficulty collected from the slider, the task performance collected as a score between 0 and 3 collected from the number of digits correctly heard and repeated, the response time score collected as the time elapsed between the appearance of the visual stimulus and the participant's clicking reaction, and the TEPR (task-evoked pupil response) collected by the Tobii 2 glasses based on the participant's eye pupil dilation as a reaction to the listening task. All participants responded to the same set of audio recordings and all steps of the procedure were identical, with the exception of individualized SNR settings for easy and hard listening conditions. The results are reported both on individual analyses of each participant and an averaging of the group. The results of the statistical tests and visual representation figures were created on the jamovi 2.5 (The jamovi project, 2024) software.

The slider score is a subjective rating of the repetition's difficulty, and the program enables the participant to give a rating on the scale of 0 to 100. For the purposes of this study, the slider score serves as a subjective rating of the perceived listening effort (Lemke & Besser, 2016) experienced by the participant. The digit performance score is the number of digits correctly guessed by the participant for each repetition, ranging from 0 to 3.

For this study, the TEPR score for each repetition is calculated based on the time window during which the participant hears the audio input and the corresponding pupil dilation. The time window of each repetition is divided into two components: a baseline window and a stimulus window.

The baseline window is the period before the participant hears the audio. During this time, the pupil is in a relaxed state, and the average pupil diameter is used as the baseline measurement. The baseline window is defined as the 2-second interval prior to the start of the audio recording. The stimulus window, defined as the duration when the participant listens to the audio, includes the length of the audio recording (ranging from 3.3 to 3.9 seconds) plus one additional second after the recording ends. Thus, the total length of the stimulus window ranges from 4.3 to 4.9 seconds.

The pupil diameter measurements from both the baseline and stimulus windows are averaged separately. The TEPR score is calculated by subtracting the baseline average from the stimulus average. This difference is then expressed as a percentage of the baseline average. This percentage represents the TEPR score for each repetition.

The response time score is calculated as follows: The time difference between the appearance of the visual stimulus (green light) and participant's time of clicking is calculated (ranging from 0.3 seconds up to a full second). A baseline score is produced by averaging the response time scores of the 20 repetitions from the baseline test. During further tests, this baseline score is subtracted from the response time for each repetition. This difference is then presented as a percentage of the baseline score, which is set as the response time score for the repetition.

All data used in the study has been subjected to Shapiro-Wilk tests (Navarro & Foxcroft, 2022) and displayed p values above 0.05, apart from the digit performance scores. These scores range from 0 to 3, and mostly concentrate on 3 or 2 correct digits, and therefore do not distribute normally. Consequently, the statistical test which did not assume normality of a data set (Wilcoxon rank sum test) was used for exploration of task performance in different difficulties.

For comparisons of the same participant's measures over easy and hard difficulties, paired samples t-tests were used. The reason for this was that paired samples t-tests are recommended to be used in cases where the same source of data is tested across different conditions. For exploration of measures' relations with each other and discovery of cooccurrences, a correlation matrices were used. The explorations are conducted both on individual participants' data, and the averaging of the five participants' data as a group. In the following discussion of analysis, this is referred to as "group average".

I started by exploring the relation between task performance and the difficulty setting. For this, I've grouped the dataset under easy and hard settings, and checked if the two difficulties' digit performance scores showed a difference.

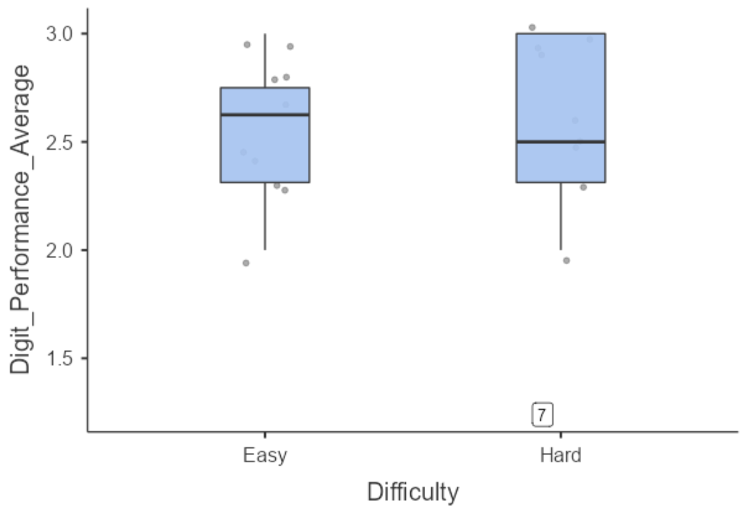

Figure 9: Box plots of digit performance of the group average of the five participants, on easy and hard task difficulties.

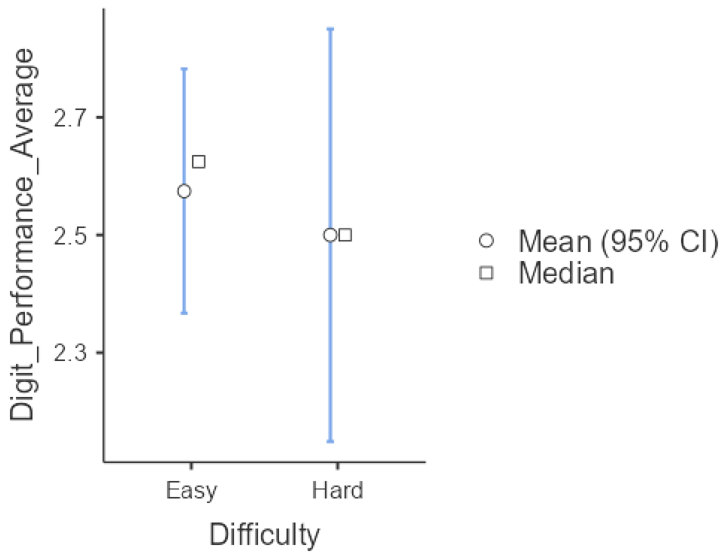

Figure 10: Descriptive plots of digit performance of the group average of the five participants, on easy and hard task difficulties.

Table 3: Paired Samples T-Test (Wilcoxon W) of Digit Performance on easy and hard task difficulties.

Participant	**Statistic**	**p**
Participant 7	21.0	0.240
Participant 8	4.00	0.773
Participant 9	20.5	0.851
Participant 10	6.00	0.766
Participant 11	0.0	0.149
Group average	18.5	1.000

Note. H_a μ Easy ≠ μ Hard

The visual representations show that the task performance on hard difficulty repetitions show a wider spread than easy difficulty. While the mean and median values for easy difficulty are higher, the paired samples Wilcoxon signed ranks test yields a statistic of 18.5 (p = 1.000), which is not statistically significant.

Next, I wanted to explore the relation between self-reported perceived effort scores and the different difficulty settings. I followed the same procedure as the previous step. Since the self-reported effort scores are collected from sliders, they are referred to as "slider" scores in the visual representations.

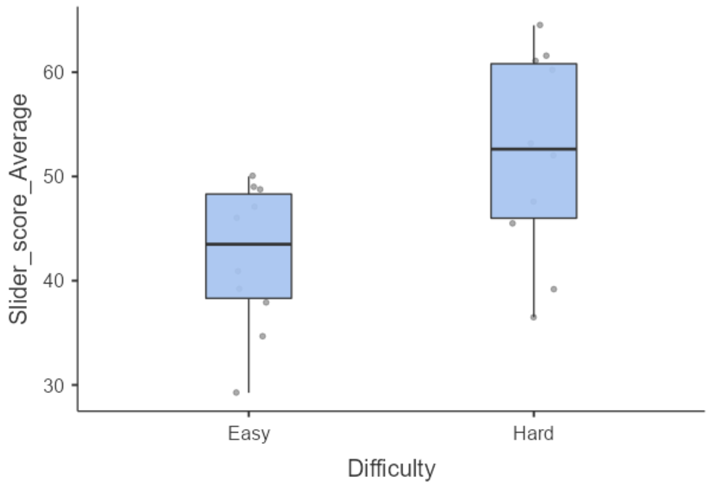

Figure 11: Box plots of self-reported effort score group average of the five participants, on easy and hard task difficulties.

Table 4: Paired Samples T-Test of self-reported effort scores, on easy and hard task difficulties.

Participant	**Statistic**	**df**	**p**
Participant 7	-2.60	9.00	0.014
Participant 8	-1.38	9.00	0.101
Participant 9	0.444	9.00	0.666
Participant 10	-1.59	9.00	0.074
Participant 11	-1.11	9.00	0.148
Group Average	-2.43	9.00	0.019

Note. H_a μ Easy < μ Hard

Here, the alternative hypothesis was that scores for the hard difficulty setting would be higher than those for the easy difficulty setting, indicating that participants reported higher effort in the harder setting. With a p-value below 0.05, this finding is moderately significant for the group average. Although only one out of five participants showed a significant increase in effort scores individually, the group average difference is significant. This likely results from the increased sample size when averaging the group's data. While the difference might not be significant for each participant, it becomes more apparent in the combined data. Therefore, it suggests that the harder difficulty elicits higher effort scores from participants.

An interesting exploration was whether the participants showed tendency to give higher self-reported effort ratings to repetitions where they made more errors. This relation is investigated with a correlation matrix. First, I applied the matrix test on the group average of the five participants.

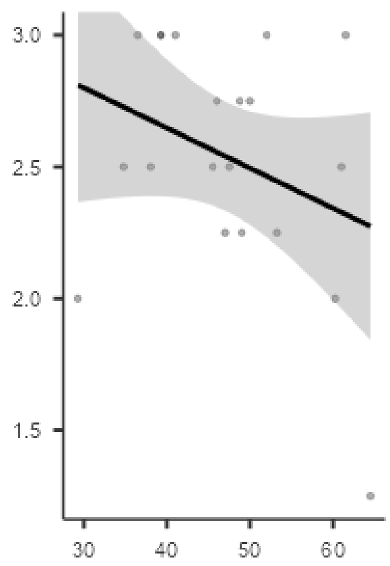

Figure 12: Correlation matrix representation of self-reported effort ratings and digit performance scores of the group of five participants. Horizontal: self-reported effort score. Vertical: digit performance score.

Table 5: Correlation matrix of self-reported effort rating and digit performance score of the group average of the five participants.

Pearson's r	-0.325
p-value	0.162

The visual representation shows a negative tendency as expected, but no significant correlation is displayed. Although when I checked the same correlation matrix on individual participants, I saw that the data from participants 8, 10 and 11 show a significant correlation.

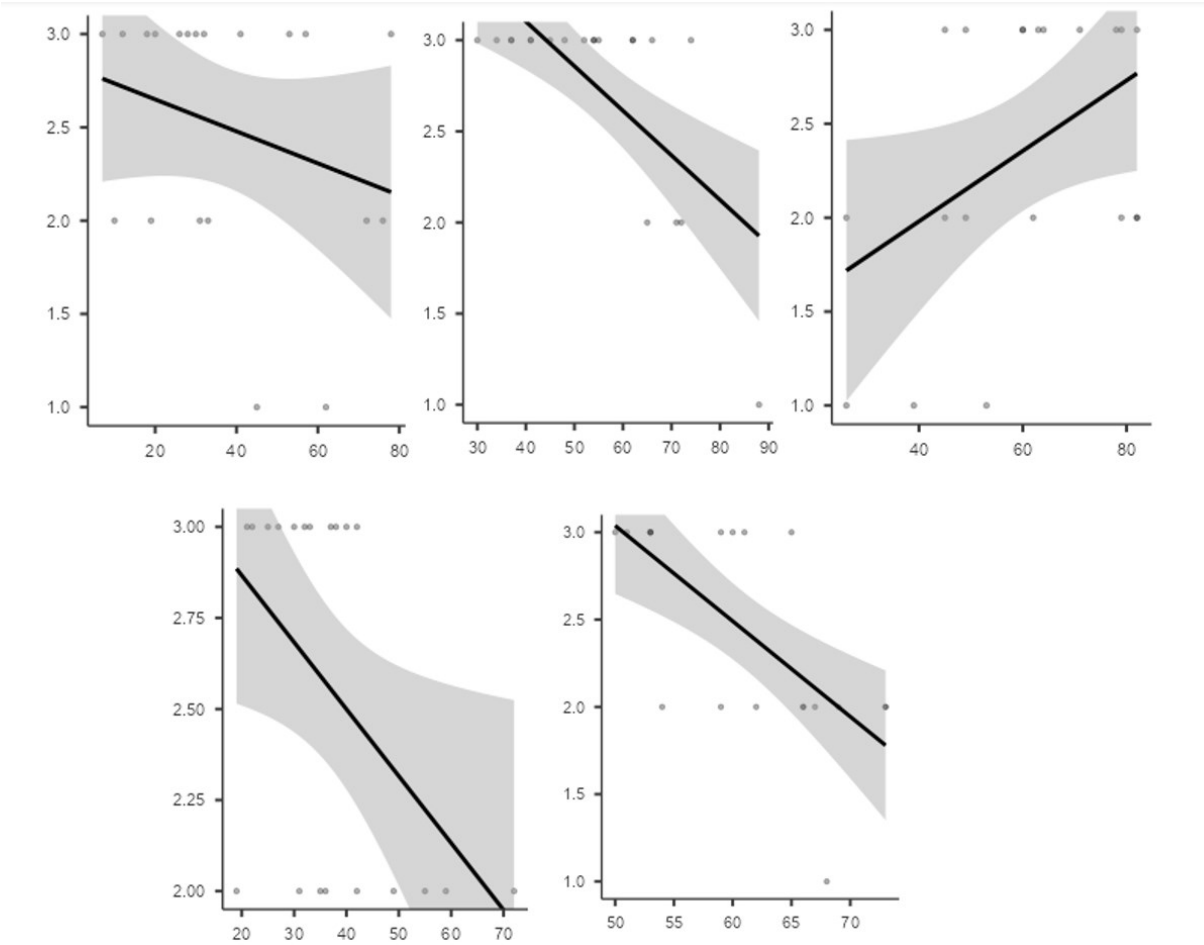

Figure 13: Correlation matrix visual representations of self-reported effort ratings and digit performance scores of participants 7, 8, 9, 10 and 11, clockwise. Horizontal: self-reported effort ratings. Vertical: digit performance scores.

Table 6: Correlation matrix of individual participants' self-reported effort ratings and digit performance scores.

Participant	Pearson’s r	p-value
Participant 7	-0.276	0.239
Participant 8	-0.691	<.001
Participant 9	0.452	0.045
Participant 10	-0.483	0.016
Participant 11	-0.668	<.001

Above are the results of the correlation matrix displaying a negative correlation between digit performance and self-reported effort scores of individual participants. The results from the individual correlation matrices show possible tendency that participants 8, 10 and 11 were more likely to report higher difficulty on repetitions where they made errors.

The analysis of TEPR and response time data did not show correlation with other data of self-reported effort ratings and the categories of easy and hard listening conditions. The relation between reaction time and task difficulty seem to mirror previous research as hard difficulty points to longer reaction time.

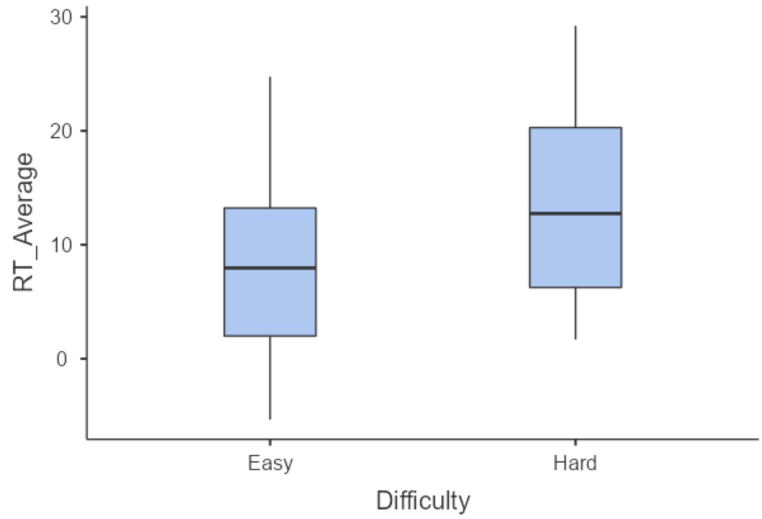

Figure 14: Box plots of RT scores of the group average of the five participants, on easy and hard task difficulties.

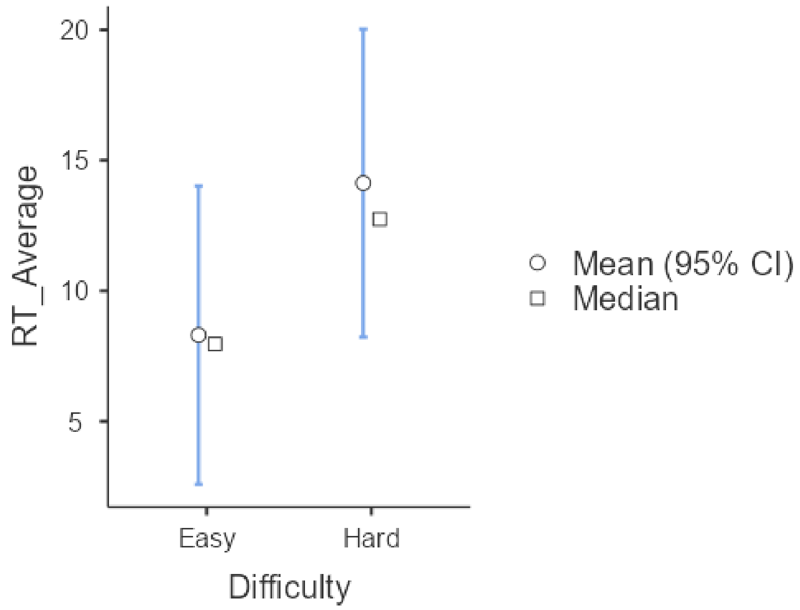

Figure 15: Descriptive plots of RT scores of the group average of the five participants, on easy and hard task difficulties.

The figures 12 and 13 are complemented by Table 5.

Table 7: Paired Samples T-Test of RT scores on easy and hard task difficulties.

Participant	**Statistic**	**df**	**p**
Participant 7	0.437	9.00	0.664
Participant 8	-0.533	9.00	0.304
Participant 9	-1.79	9.00	0.053
Participant 10	-1.90	9.00	0.045
Participant 11	-0.721	9.00	0.244
Group Average	-1.70	9.00	0.062

Note. H_a μ Easy < μ Hard

Although the hard difficulty RT score tends to be higher than easy difficulty for participants, the group average RT score does not show statistically significant difference with a p value of 0.062. The only participant which shows significant tendency to have a longer response time in hard difficulty is Participant 10 with a p value of 0.045. Refer to 4.4.3 *Reliability* for more information on TEPR and RT measurements.

In summary, the listening effort study has reported these results:

1. The digit performance, while having a wider distribution in hard difficulty, is not significantly correlated to difficulty. **Easy and hard block repetitions show mostly similar correct digit rates.**

2. The self-reported effort scores of participants show difference in regard to difficulty. **Hard block repetitions are rated to be more difficult to listen and comprehend accurately.**

3. By some individual datasets, slider scores show significant negative correlation to digit performance rates. **Participants tend to rate repetitions to be more difficult if they have made an error.**

6. Results and Discussion

6.1 Literature Review Discussion

The first research question for this study being “What are the factors influencing the characteristics of mainstream foreign language lessons for hard-of-hearing children, as reported in previous literature?”, the literature review was conducted to explore previous studies on the subject in order to gain a wider perspective. From the overview offered by the literature review, the process of which led to the conceptualization of this study, it is possible to list research-based facts on the subject of language education for hard-of-hearing children and how listening effort might influence it. It is known that most hard-of-hearing children receive early intervention and with a correlation of how early they were implanted, they can continue their language development appropriate to their age (Cuda et al., 2014; Shojaei et al., 2016; Ruben, 2018). As a result, most follow mainstream classes alongside their normal hearing peers. We know that despite the technological advancements made in the last 20 years, hard-of-hearing students form a very heterogenous group depending on the needs of the individual child, their technological devices and the environments they are in, and that individualized attention and support form the backbone of their future academic success (FEADPA, 2023; Archbold & Mayer, 2012; Nussbaum, 2018). Finally, we know that a high buildup of cognitive load from extrinsic factors can overload the working memory, which will impede task performance and learning (see section 5.1.3 *Cognitive Load Theory and Listening Effort*). Coupled with the added effect of listening effort on listening tasks, the experience of hard-of-hearing children during foreign language classes might suffer. A culmination of the cumulative summary of previous studies offers the following factors as influences on mainstream early language education of hard-of-hearing students:

1. **Academic success is achievable:** Through early intervention and adequate support provided by both medical methods and educational planning, hard-of-hearing children can thrive in mainstream education, achieving academic success in general curriculum and foreign language courses.
2. **The experience of each child is different:** Hard-of-hearing children are a heterogenous group with high individual variance, therefore hard-of-hearing children may have varying levels of accustoming to mainstream education, familiarity with sign language, motivation to learn a foreign language, academic success and support provided by the three-tiered support system.
3. **Isolation and frustration are common situations:** Hard-of-hearing students tend to experience social isolation and underdeveloped social language skills; early implanted children tend to resume their age-appropriate development and are able to find success in foreign language outcomes.
4. **Heavier load on working memory:** Hard-of-hearing children may experience more cognitive load on working memory during listening tasks therefore struggle in certain foreign language classroom practices, especially in complex listening tasks and listening comprehension. This can result in tiredness and decreased performance.
5. **Classroom practices may help:** The differentiation principles and teachings from cognitive load theory research can be used to develop modifications to lessons which can decrease the load on the working memory of hard-of-hearing students.
6. **Cooperation is important:** The cooperation between the teacher, medical specialists, school support staff and the parents is important to support hard-of-hearing students in mainstream education.

6.2 Listening Effort Study Discussion

One of the goals of this study was to explore the relations of different parameters during an effortful listening task in order to better understand how an adverse hearing condition can affect listening. The second research question, “How do self-reported listening effort, task performance and task difficulty relate in a listening task under difficult hearing conditions?”, was formulated to complement this research goal, and the listening effort study was conducted to explore possible answers. As stated in the section 3.2 “*Prediction of Results*”, the expected outcomes are that, based on the increase in hearing difficulty, the error rate and self-reported effort rate will also increase. The study pointed to three main findings, which will be discussed in order.

6.2.1 Task Performance and Difficulty

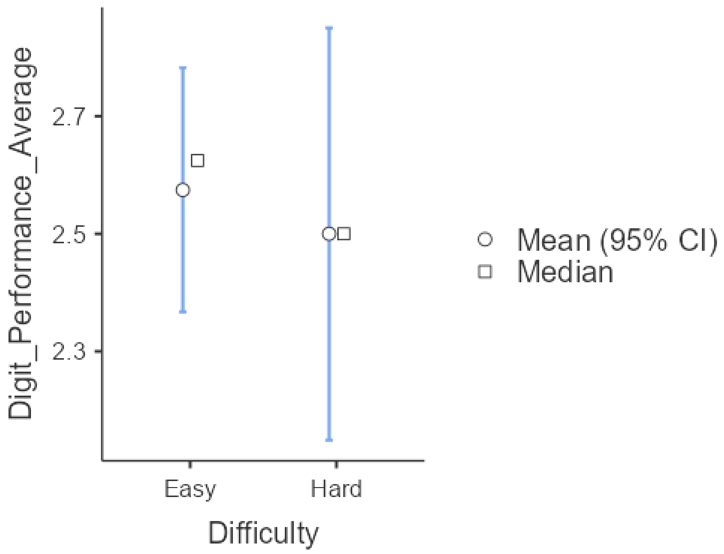

Figure 16: Descriptive plots of task performance group average on easy and hard task difficulties.

Table 8: Paired Samples T-Test (Wilcoxon W) results of task performance in easy and hard conditions.

Participant	**Statistic**	**p**
Participant 7	21.0	0.240
Participant 8	4.00	0.773
Participant 9	20.5	0.851
Participant 10	6.00	0.766
Participant 11	0.0	0.149
Group average	18.5	1.000

Note. H_a μ Easy ≠ μ Hard

The paired samples Wilcoxon signed ranks test of task performance (the number of correct digits repeated) returned no results to indicate a relation. The only finding was that the hard difficulty conditions resulted in a larger spread of correctly guessed digit rate, but there was no finding that the task performance in hard difficulty resulted in statistically suffered from more errors. The expected results were that there could be an observable difference in task performance between the two difficulty settings, although this is not reflected in the findings. For example, in McGarrigle et al.'s (2020) study, the results of Figure 17 can be seen. The easy difficulty "percentage correct" rate (over 90%) is higher than the hard difficulty (%60-75). A possible reason for this discrepancy might be that in McGarrigle et al. (2020), the easy difficulty was defined as 5 dB higher that the hard (L84) difficulty, which resulted in a fixed dB difference of 5. In this study, the easy and hard task difficulties were defined as L50 and L90, which resulted in dB difference of less than 5 for the two difficulties. Also, in most listening effort studies, the listening task is composed of sentence-level information retrieval and creating input based on the audio. In this study, the participants only listen to three digits and repeat them. It can be hypothesized that as task complexity increases, the difference in task performance might also be more observable.

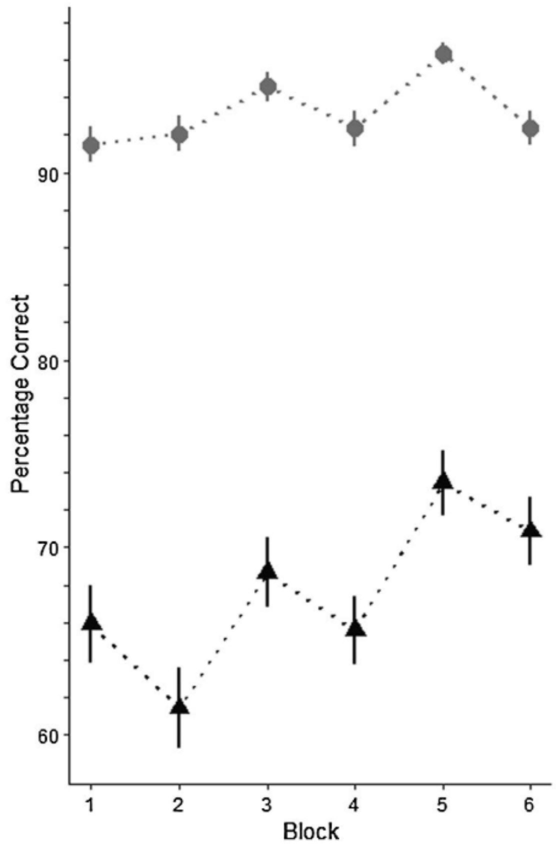

Figure 17: Figure from McGarrigle et al. (2022), depicting a higher percentage of correct answers in easy difficulty than hard. Dots are data points from easy repetitions, and triangles are from hard repetitions.

Another factor for the lack of statistical difference in task performance among easy and hard task difficulties can also be that while the hearing difficulty is increases, the cognitive resources of the participant are being allocated to the task to complement the increased difficulty, resulting in a similar success rate (Hunter, 2021; Hunter & Pisoni, 2018). According to previous research (Pals et al., 2013), a focused allocation such as this would create a burdening demand on working memory, effects of which could be observed by dual-task measures. The response time (RT) measurement of the listening effort test can provide insight on this theory.

To observe if the increased allocation of cognitive resources during the hard difficulty affects the participants' performance of secondary tasks, the RT scores of the two difficulties can be compared. The following are the paired samples t-test results of individual participants and the average of the group.

Table 9: Paired Samples T-test of RT scores in easy and hard task difficulties.

Participant	Statistic	df	p
Participant 7	0.437	9.00	0.664
Participant 8	-0.533	9.00	0.304
Participant 9	-1.79	9.00	0.053
Participant 10	-1.90	9.00	0.045
Participant 11	-0.721	9.00	0.244
Group Average	-1.70	9.00	0.062

Note. H_a μ Easy < μ Hard

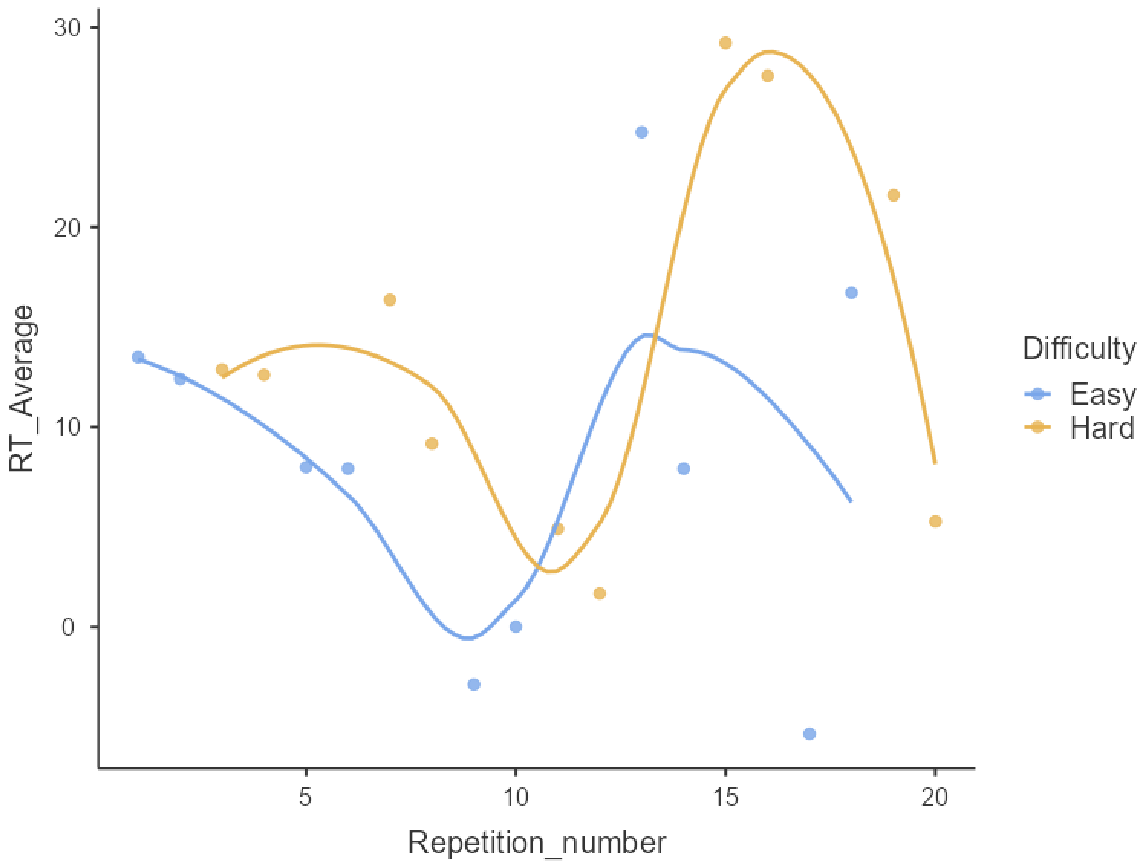

Figure 18: Scatterplot of group average RT scores with repetition numbers, in easy and hard task difficulties.

While Participant 10’s test result yields a p value of 0.045, the other participants’ results, and the group average do not return a significant difference between the RT scores of easy and hard difficulties. While the visual representation helps observe that the hard difficulty RT scores have a higher range (see also Figures 12 and 13); the low sample size and reliability issues with RT measurements (4.4.3. *Reliability*) require caution in interpreting this as a definite finding.

6.2.2 Self-Reported Effort and Difficulty

The exploration of the relation between self-reported effort and difficulty leads to a significant finding of the study.

Table 10: Paired Samples T-Test of self-reported effort in easy and hard task difficulties.

Participant	**Statistic**	**df**	**p**
Participant 7	-2.60	9.00	0.014
Participant 8	-1.38	9.00	0.101
Participant 9	0.444	9.00	0.666
Participant 10	-1.59	9.00	0.074
Participant 11	-1.11	9.00	0.148
Group Average	-2.43	9.0	0.019

Note. H_a μ Easy < μ Hard

The t-test applied to the findings show that increased hearing difficulty shows a statistically significant cooccurrence with an increased perceived effort. Combined with the finding mentioned before that the easy and hard difficulties show no significant difference in task performance, an explanation of the results might be as such: the overloading of the working memory to compensate for the difficult hearing condition might result in no significant drop in task performance, but the overexertion might give the listener a sense of discomfort, possibly leading to faster accumulation of fatigue. McGarrigle et al. (2020) shows that the participants report an increased level of fatigue as the study continues, and Shields et al. (2022) reports cochlear implanted patients, who experience adverse hearing conditions, report higher measures of listening effort.

6.2.3. Self-reported Effort and Task Performance

The final finding is in regard to the exploration of self-reported effort and digit performance.

Table 11: Correlation Matrix of self-reported effort ratings and digit performance scores.

Participant	**Pearson's r**	**p-value**
Participant 7	-0.276	0.239
Participant 8	-0.691	<.001
Participant 9	0.452	0.045
Participant 10	-0.483	0.016
Participant 11	-0.668	<.001
Group Average	-0.325	0.162

While the averaged statistic is not significant, the individual results of participants 8, 10 and 11 show a negative correlation of self-reported effort scores and correct digit rates. This is to indicate a cooccurrence of increased perception of effort and errors made in the listening task on an individual basis.

Overall, the findings from the listening effort study show that listeners who experience more difficult hearing conditions, while being able to perform at an adequate rate, can suffer more feelings of overexertion. Superficial findings suggest that this might lead to decreased performance in secondary tasks. On an individual basis, listeners may tend to feel more effort in situations where they make mistakes.

7. Conclusion

7.1 Summary

The aims of this study were to provide an overview of hard-of-hearing students' experiences in mainstream early language education, and to fill the gap in previous research by exploring how the increased listening effort might affect their performance in their language classrooms, particularly with listening tasks. A mixed method approach, including a quantitative listening effort study and a qualitative narrative literature review, was conducted to achieve these goals. The discussion regarding the correspondence of formulated research questions with the resulting insights from the methods of this study is aimed to offer a wide comprehension of the phenomenon being explored.

In response to the first research question "What is the effect of listening effort on the mainstream early language education of hard-of-hearing students?", the findings of the literature review first highlight the documented experiences of hard-of-hearing children within mainstream education settings. These students exhibit varying characteristics in learning and demands of differentiation. Social isolation and frustration are common challenges faced by hard-of-hearing students, while those who receive early cochlear implants tend to show improved language development and foreign language outcomes. Furthermore, hard-of-hearing children may experience heightened cognitive load on working memory and increased listening effort during listening tasks, leading to difficulties in complex listening tasks and comprehension. To explore the factors which listening effort can influence during a listening task, the findings of the listening effort test can be referenced.

The second research question "How do self-reported listening effort, task performance and task difficulty relate in a listening task under difficult hearing conditions?" is investigated by the listening effort study. The listening effort tests demonstrate the differences in easy and adverse conditions of hearing during a listening task. In the harder difficulty, the listeners are found to report perceiving a higher amount of effort to listen, even if they are able to maintain a satisfactory level of performance. The participants report heightened sensations of overexertion, potentially leading to diminished performance in concurrent tasks. Some participants' results also show a correlation between increased perception of effort and instances of error during task execution. This suggests that on a personal level, individuals may be more prone to perceive heightened effort in situations where errors occur.

The effect of listening effort on the mainstream early language education of hard-of-hearing students can be explored by combining the qualitative and quantitative findings. The cognitive load theory asserts that working memory capacity spent on overcoming external factors deplete the capacity that can otherwise be allocated to comprehension and learning. Hard-of-hearing students experience such external factors during listening tasks in language classes, due to the effects of listening effort. The quantitative findings suggest that even when hard-of-hearing students might perform at a similar level to normal hearing students, they may experience more feelings of overexertion and might be overloaded during tasks with high complexity, or multitasking activities.

The findings of this study show that hard-of-hearing children in mainstream education face diverse challenges. Despite these challenges, hard-of-hearing students are capable of academic success and attaining foreign language competency through mainstreamed early language education. Support offered through both educational planning and the initiative used by teachers to differentiate content can assist this premise. In this endeavour, the individual differences and needs of the students must be considered. Social isolation is commonly experienced, which underscores the importance of collaboration among educators, specialists, and parents. The strategies employed to improve performance can benefit from guiding principles derived from cognitive load theory and an understanding of the effects of listening effort. It is hoped that these insights can provide an outlook for future efforts to improve the quality of early language education for hard-of-hearing students.

7.2 Limitations

The primary constraint of the study stems from its reliance on hard-of-hearing students as the focus population, yet the inability to establish direct contact with such students hinders the depth of insights that could have been gleaned. Furthermore, in the quantitative aspect of the study, the participant group consisted of only five individuals, a notably smaller sample size compared to established studies in the field, which typically employ power analysis methods to determine sample size adequacy. This was due to the constraint that the equipment used for the study was shared among other research groups and projects, and the time required to form a larger participant pool while also retaining the equipment was not available. Consequently, the conclusions drawn in this study were primarily based on insights already demonstrated in previous research. Additionally, the use of pupillometry as a measure proved to be subject to low test-retest reliability, failing to capture correlations observed in some prior studies.

7.3 Future Research Possibilities

In future research involving listening effort and hard-of-hearing education, the primary recommendation offered by this study is to establish contact with the target population of hard-of-hearing students via their parents and teachers. An interview with parents and teachers might reveal more valuable insights directly from informants who experience the subject of the study. As such an interview was planned for this study, I've prepared a set of guiding questions for teachers of hard-of-hearing students, which can be found at the appendix section. Another prospect would be to prepare a listening effort study protocol similar to the one used in this study, and test with control and test groups of normal hearing and hard-of-hearing children in order to gain further insight into the factors by which listening effort affects a listening task.

References

ADCET. (2015, February). Inclusive teaching: Deaf and hard of hearing. Australian Disability Clearinghouse on Education and Training. Retrieved April 22, 2024, from https://www.adcet.edu.au/inclusive-teaching/specific-disabilities/deaf-hearing-impaired

Aguilar-Mediavilla, E., Pérez-Pereira, M., Serrat-Sellabona, E., & Adrover-Roig, D. (2022). Introduction to Language Development in Children: Description to Detect and Prevent Language Difficulties. Children (Basel, Switzerland), 9(3), 412. https://doi.org/10.3390/children9030412

Allen, T. E., & Morere, D. A. (2020). Early visual language skills affect the trajectory of literacy gains over a three-year period of time for preschool aged deaf children who experience signing in the home. PloS one, 15(2), e0229591. https://doi.org/10.1371/journal.pone.0229591

Archbold, S., & Mayer, C. (2012). Deaf Education: The Impact of Cochlear Implantation? Deafness & Education International, 14(1), 2–15. https://doi-org.ezproxy.uef.fi:2443/10.1179/1557069X12Y.0000000003

ASHA. (2015). Effects of hearing loss on development. American Speech-Language-Hearing Association. https://www.asha.org/public/hearing/effects-of-hearing-loss-on-development/

ASHA. (2023, November). ASHA's developmental milestones: Birth to 5 years. American Speech-Language-Hearing Association. https://www.asha.org/public/developmental-milestones

Bienvenu, MJ and Meehan, Jessica (2022) "A Tribute to Marie Jean Philip," Society for American Sign Language Journal: Vol. 6: No. 2, Article 4. Available at: https://tigerprints.clemson.edu/saslj/vol6/iss2/4

Blair, J. C., & Langan, L. (2000). Can you hear me?. A longitudinal study of hearing aid monitoring in the classroom. Journal of Educational Audiology, 8, 34-36.

Briscoe, J., Bishop, D. V., & Norbury, C. F. (2001). Phonological processing, language, and literacy: a comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. Journal of child psychology and psychiatry, and allied disciplines, 42(3), 329–340.

CESE. (2017, September). Cognitive load theory: Research that teachers really need to understand. Centre for Education Statistics and Evaluation. https://education.nsw.gov.au/about-us/educational-data/cese/publications/literature-reviews/cognitive-load-theory

Cleveland Clinic. (2023, March). Cochlear implants. Cleveland Clinic. Retrieved April 18, 2024, from https://my.clevelandclinic.org/health/treatments/4806-cochlear-implants

Cohen, L., Manion, L., & Morrison, K. (2018). Mixed methods research. In Research methods in education (8th ed., pp. 31–50). essay, Routledge.

Colenbrander, D., Nickels, L., & Kohnen, S. (2011). Nonword reading tests: A review of the available resources. Australasian Journal of Special Education, 35(2), 137–172. https://doi.org/10.1375/ajse.35.2.137

Creswell, J. W., & L., P. C. V. (2018). Designing and conducting mixed methods research. SAGE.
Cuda, D., Murri, A., Guerzoni, L., Fabrizi, E., & Mariani, V. (2014). Pre-school children have better spoken language when early implanted. International journal of pediatric otorhinolaryngology, 78(8), 1327–1331. https://doi.org/10.1016/j.ijporl.2014.05.021
de Jong, T. (2009). Cognitive load theory, educational research, and instructional design: Some Food For Thought. Instructional Science, 38(2), 105–134. https://doi.org/10.1007/s11251-009-9110-0
Deunk, M. I., Smale-Jacobse, A. E., de Boer, H., Doolaard, S., & Bosker, R. J. (2018). Effective differentiation Practices:A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education. Educational Research Review, 24, 31–54. https://doi.org/10.1016/j.edurev.2018.02.002
DO-IT. (2022). | How are the terms deaf, deafened, hard of hearing, and hearing impaired typically used? Date updated April 24, 2022. Retrieved May 2, 2023, from https://www.washington.edu/doit/how-are-terms-deaf-deafened-hard-hearing-and-hearing-impaired-typically-used
Everhardt, M. K., Jung, D. E., Stiensma, B., Lowie, W., Başkent, D., & Sarampalis, A. (2023). Foreign Language Acquisition in Adolescent Cochlear Implant Users. Ear & Hearing, 45(1), 174–185. https://doi.org/10.1097/aud.0000000000001410
FEADPA. (2023, April 28). Comm4child project fact sheet 2023. The European Federation of Associations of Teachers of the Deaf . https://www.feapda.eu/projects/comm4child-project-fact-sheet-2023
Felman, A. (2023). Deafness and hearing loss: Causes, symptoms, and treatments. Medical News Today. Retrieved April 17, 2024, from https://www.medicalnewstoday.com/articles/249285
Fetters, M. D., & Freshwater, D. (2015). The 1 + 1 = 3 Integration Challenge. Journal of Mixed Methods Research, 9(2), 115–117. https://doi.org/10.1177/1558689815581222
Finnish National Board on Research Integrity TENK. (2023) The Finnish Code of Conduct for Research Integrity and Procedures for Handling Alleged Violations of Research Integrity in Finland. Guideline of the Finnish National Board on Research Integrity TENK 2023. Retrieved May 16, 2024, from https://tenk.fi/sites/default/files/2023-11/RI_Guidelines_2023.pdf
Friesen, L. (2019, October 30). What does the world sound like through a cochlear implant?. Cochlear Implant Brain and Behavior Lab. Retrieved April 18, 2024, from https://cochlearimplant.lab.uconn.edu/cochlear-implant-information/sounds/
Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The Children's Test of Nonword Repetition: a test of phonological working memory. Memory (Hove, England), 2(2), 103–127. https://doi.org/10.1080/09658219408258940
Geers, A. E., Mitchell, C. M., Warner-Czyz, A., Wang, N. Y., Eisenberg, L. S., & CDaCI Investigative Team (2017). Early Sign Language Exposure and Cochlear Implantation Benefits. Pediatrics, 140(1), e20163489. https://doi.org/10.1542/peds.2016-3489

Geers, A. E., Nicholas, J. G., & Sedey, A. L. (2003). Language skills of children with early cochlear implantation. Ear and hearing, 24(1 Suppl), 46S–58S. https://doi.org/10.1097/01.AUD.0000051689.57380.1B

Giuliani, N. P., Brown, C. J., & Wu, Y.-H. (2021). Comparisons of the Sensitivity and Reliability of Multiple Measures of Listening Effort. Ear & Hearing, 42(2), 465–474. https://doi.org/10.1097/AUD.0000000000000950

Goldin-Meadow, Susan; Mayberry, Rachel I. (November 2001). "How Do Profoundly Deaf Children Learn to Read?". Learning Disabilities Research and Practice. 16 (4): 222–229. doi:10.1111/0938-8982.00022.

Granholm, E., & Steinhauer, S. R. (2004). Pupillometric measures of cognitive and emotional processes. International Journal of Psychophysiology, 52(1), 1–6. https://doi.org/10.1016/j.ijpsycho.2003.12.001

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: a test of the critical-period hypothesis for second-language acquisition. Psychological science, 14(1), 31–38. https://doi.org/10.1111/1467-9280.01415

Handley, C. (2021, August 2). Cognitive load theory and the differences between experts and novices: What chess tells us about teaching foreign languages. Mind Brain Education. https://www.mindbrained.org/2021/08/cognitive-load-theory-and-the-differences-between-experts-and-novices-what-chess-tells-us-about-teaching-foreign-languages/

Hanham, J., Castro-Alonso, J. C., & Chen, O. (2023). Integrating cognitive load theory with other theories, within and beyond educational psychology. The British journal of educational psychology, 93 Suppl 2, 239–250. https://doi.org/10.1111/bjep.12612

Harvard Catalyst. (2022). Mixed methods research. Mixed Methods Research - Harvard Catalyst. https://catalyst.harvard.edu/community-engagement/mmr/

Hunter C. R. (2021). Dual-Task Accuracy and Response Time Index Effects of Spoken Sentence Predictability and Cognitive Load on Listening Effort. Trends in hearing, 25, 23312165211018092. https://doi.org/10.1177/23312165211018092

Hunter, C. R., & Pisoni, D. B. (2018). Extrinsic Cognitive Load Impairs Spoken Word Recognition in High- and Low-Predictability Sentences. Ear and hearing, 39(2), 378–389. https://doi.org/10.1097/AUD.0000000000000493

ICED. (2018, June 11). 21st International Congress on the Education of the Deaf (ICED) in July 2010 in Vancouver, Canada. WFD. https://wfdeaf.org/news/21st-international-congress-on-the-education-of-the-deaf-iced-in-july-2010-in-vancouver-canada/

Jhangiani, R. S., & Price, P. C. (2017, August 21). 4.2 Reliability And Validity Of Measurement. Research methods in psychology. https://opentext.wsu.edu/carriecuttler/chapter/reliability-and-validity-of-measurement/

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. Journal of Mixed Methods Research, 1(2), 112–133. https://doi.org/10.1177/1558689806298224

Koelewijn, T., Zekveld, A. A., Festen, J. M., & Kramer, S. E. (2012). Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. Ear and hearing, 33(2), 291–300. https://doi.org/10.1097/AUD.0b013e3182310019

Kontra, E. H. (2013). Language learning against the odds: retrospective accounts by Deaf adults. In English as a foreign language for deaf and hard-of-hearing persons in Europe (pp. 93–109). essay, Wydawnictwo KUL.

Lang, H. G. (2011). Perspectives on the History of Deaf Education. In The Oxford Handbook of Deaf Studies, Language, and Education (2nd ed., Vol. 1, pp. 7–17). essay, Oxford University Press.

Lemke, U., & Besser, J. (2016). Cognitive Load and Listening Effort: Concepts and Age-Related Considerations. Ear and hearing, 37 Suppl 1, 77S–84S. https://doi.org/10.1097/AUD.0000000000000304

Lonka, E. (2014). From Silence To Sounds : Outcomes Of Cochlear Implantation In Finnish Adults And Children (dissertation). University of Helsinki, Helsinki. Retrieved March 12, 2024, from https://helda.helsinki.fi/bitstreams/fbfdce30-2518-4d94-8d00-f2cd48f65bf6/download

Lynch, T. (2015). International students' perceptions of university lectures in English. ISEJ – International Student Experience Journal, 3(1). Retrieved 8 May 2024 from, https://cdn.southampton.ac.uk/assets/imported/transforms/content-block/UsefulDownloads_Download/76D63415CCE1466098F6BF6CD3467F54/Lynch%20(2015)%20International%20students'%20perceptions%20of%20lecturers%20in%20English.pdf

Manley, J. (2020, October 12). 5 language activities for children with hearing loss. Central Institute for the Deaf Professional Development. Retrieved April 22, 2024, from https://professionals.cid.edu/five-language-activities-children-hearing-loss-age-language-level/

Marchman, V. A., & Dale, P. S. (2023). The MacArthur-Bates Communicative Development Inventories: updates from the CDI Advisory Board. Frontiers in psychology, 14, 1170303. https://doi.org/10.3389/fpsyg.2023.1170303

Marsella, P., Scorpecci, A., Cartocci, G., Giannantonio, S., Maglione, A. G., Venuti, I., Brizi, A., & Babiloni, F. (2017). EEG activity as an objective measure of cognitive load during effortful listening: A study on pediatric subjects with bilateral, asymmetric sensorineural hearing loss. International Journal of Pediatric Otorhinolaryngology, 99, 1–7. https://doi.org/10.1016/j.ijporl.2017.05.006

Mathews, E. S. (2017). Chapter 2: Mainstreaming Deaf Education. In Language, Power, and Resistance: Mainstreaming Deaf Education (pp. 18–41). essay, Gallaudet University Press.

Mathôt S. (2018). Pupillometry: Psychology, Physiology, and Function. Journal of cognition, 1(1), 16. https://doi.org/10.5334/joc.18

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., & Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in hearing special interest group 'white paper.' International Journal of Audiology, 53(7), 433–445. https://doi.org/10.3109/14992027.2014.890296

McGarrigle, R., Rakusen, L., & Mattys, S. (2020). Effortful listening under the microscope: Examining relations between pupillometric and subjective markers of effort and tiredness from listening. Psychophysiology, 58(1). https://doi.org/10.1111/psyp.13703

Moeller, M. P., McCleary, E., Putman, C., Tyler-Krings, A., Hoover, B., & Stelmachowicz, P. (2010). Longitudinal development of phonology and morphology in children with late-identified mild-moderate sensorineural hearing loss. Ear and hearing, 31(5), 625–635. https://doi.org/10.1097/AUD.0b013e3181df5cc2

Moores, D.F. (2010). Partners in Progress: The 21st International Congress on Education of the Deaf and the Repudiation of the 1880 Congress of Milan. American Annals of the Deaf 155(3), 309-310. https://doi.org/10.1353/aad.2010.0016.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. Journal of Educational Psychology, 87(2), 319–334. https://doi.org/10.1037/0022-0663.87.2.319

NAD. (2020, August 4). Position statement on Dual Language Development and use in the educational environment. National Association of the Deaf. https://www.nad.org/about-us/position-statements/position-statement-on-dual-language-development-and-use-in-the-educational-environment/

Nadege, C., Garnault, V., Ferlicoq, L., Derumeaux-Burel, H., Bongard, V., Deguine, O., Fraysse, B., & Molinier, L. (2011). The cost of cochlear implantation: a review of methodological considerations. International Journal of Otolaryngology, 1–13. https://doi.org/10.1155/2011/210838

Nagels, L., Bastiaanse, R., Başkent, D., & Wagner, A. (2020). Individual differences in lexical access among cochlear implant users. Journal of Speech, Language, and Hearing Research, 63(1), 286–304. https://doi.org/10.1044/2019_jslhr-19-00192

Navarro DJ and Foxcroft DR (2022). learning statistics with jamovi: a tutorial for psychology students and other beginners. (Version 0.75). DOI: 10.24384/hgc3-7p15

NIDCD. (2010, September). Speech and language developmental milestones. National Institute of Deafness and Other Communication Disorders. https://www.nidcd.nih.gov/health/speech-and-language

Onwuegbuzie, A. J., & Hitchcock, J. H. (2019). Toward a fully integrated approach to mixed methods research via the 1 + 1 = 1 integration approach: Mixed Research 2.0. INTERNATIONAL JOURNAL OF MULTIPLE RESEARCH APPROACHES, 11(1), 7–28. https://doi.org/10.29034/ijmra.v11n1editorial2

Ouwehand, K., Kroef, A. van, Wong, J., & Paas, F. (2021). Measuring cognitive load: Are there more valid alternatives to likert rating scales? Frontiers in Education, 6. https://doi.org/10.3389/feduc.2021.702616

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. Journal of Educational Psychology, 84(4), 429–434. https://doi.org/10.1037//0022-0663.84.4.429

Pals, C., Sarampalis, A., & Başkent, D. (2013). Listening effort with cochlear implant simulations. Journal of speech, language, and hearing research : JSLHR, 56(4), 1075–1084. https://doi.org/10.1044/1092-4388(2012/12-0074)

Pals, C., Sarampalis, A., van Rijn, H., & Başkent, D. (2015). Validation of a simple response-time measure of listening effort. The Journal of the Acoustical Society of America, 138(3). https://doi.org/10.1121/1.4929614

Paré, G., & Kitsiou, S. (2017). Chapter 9 Methods for Literature Reviews. In F. Lau, & C. Kuziemsky (Eds.), Handbook of eHealth Evaluation: An Evidence-Based Approach.

University of Victoria. Retrieved 5 May 2024 from, https://www.ncbi.nlm.nih.gov/books/NBK48158
Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C. L., Naylor, G., Phillips, N. A., Richter, M., Rudner, M., Sommers, M. S., Tremblay, K. L., & Wingfield, A. (2016). Hearing Impairment and Cognitive Energy: The Framework for Understanding Effortful Listening (FUEL). Ear and hearing, 37 Suppl 1, 5S–27S. https://doi.org/10.1097/AUD.0000000000000312
Pritchard, P. (2013). Teaching of English to Deaf and severely hard-of-hearing pupils in Norway . In English as a foreign language for deaf and hard-of-hearing persons in Europe (pp. 113–134). essay, Wydawnictwo KUL.
Riva, M. M. (2023). What are decibels and how can they affect hearing health?. NCOA Adviser. Retrieved April 18, 2024, from https://www.ncoa.org/adviser/hearing-aids/decibel-levels/
Rodrigues, A., & Befi-Lopes, D. M. (2009). Phonological working memory and its relationship with language development in children. Pro-fono : revista de atualizacao cientifica, 21(1), 63–68. https://doi.org/10.1590/s0104-56872009000100011
Roiha, A. S. (2014). Teachers' views on differentiation in content and language integrated learning (CLIL): Perceptions, practices and challenges. Language and Education, 28(1), 1–18. https://doi.org/10.1080/09500782.2012.748061
Ruben R. J. (2018). Language development in the pediatric cochlear implant patient. Laryngoscope investigative otolaryngology, 3(3), 209–213. https://doi.org/10.1002/lio2.156
Saksida, A., Ghiselli, S., Bembich, S., Scorpecci, A., Giannantonio, S., Resca, A., Marsella, P., & Orzan, E. (2021). Interdisciplinary Approaches to the Study of Listening Effort in Young Children with Cochlear Implants. Audiology Research, 12(1), 1–9. https://doi.org/10.3390/audiolres12010001
Schoonenboom, J., & Johnson, R. B. (2017). How to Construct a Mixed Methods Research Design. Kolner Zeitschrift fur Soziologie und Sozialpsychologie, 69(Suppl 2), 107–131. https://doi.org/10.1007/s11577-017-0454-1
Sharma, S. D., Cushing, S. L., Papsin, B. C., & Gordon, K. A. (2020). Hearing and speech benefits of cochlear implantation in children: A review of the literature. International journal of pediatric otorhinolaryngology, 133, 109984. https://doi.org/10.1016/j.ijporl.2020.109984
Shields, C., Willis, H., Nichani, J., Sladen, M., & Kluk-de Kort, K. (2022). Listening effort: WHAT is it, HOW is it measured and WHY is it important? Cochlear Implants International, 23(2), 114–117. https://doi.org/10.1080/14670100.2021.1992941
Shojaei, E., Jafari, Z., & Gholami, M. (2016). Effect of Early Intervention on Language Development in Hearing-Impaired Children. Iranian journal of otorhinolaryngology, 28(84), 13–21.
Singleton, D. (2001). Age and second language acquisition. Annual Review of Applied Linguistics, 21, 77-89. DOI:10.1017/S0267190501000058
Soleymani, Z., Mahmoodabadi, N., & Nouri, M. M. (2016). Language skills and phonological awareness in children with cochlear implants and normal

hearing. International journal of pediatric otorhinolaryngology, 83, 16–21. https://doi.org/10.1016/j.ijporl.2016.01.013
Sundqvist C., Björk-Åman C. & Ström K. (2019) The three-tiered support system and the special education teachers' role in Swedish-speaking schools in Finland, European Journal of Special Needs Education, 34:5, 601-616, DOI: 10.1080/08856257.2019.1572094
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7
The jamovi project (2024). jamovi (Version 2.5) [Computer Software]. Retrieved from https://www.jamovi.org
Torre, A. A., Bento, R. F., Tsuji, R. K., Brito, R. de, & Bittencourt, A. G. (2012). Prelingual deafness: Benefits from cochlear implants versus conventional hearing aids. International Archives of Otorhinolaryngology, 16(03), 387–390. https://doi.org/10.7162/s1809-97772012000300014
U.S. Department of Health and Human Services. (2022). Hearing aids. National Institute of Deafness and Other Communication Disorders. Retrieved April 18, 2024, from https://www.nidcd.nih.gov/health/hearing-aids
U.S. Department of Health and Human Services. (2024). Cochlear implants. National Institute of Deafness and Other Communication Disorders. Retrieved April 18, 2024, from https://www.nidcd.nih.gov/health/cochlear-implants
UEF. (2023, November 1). Guidance on research ethics. UEF Kamu. https://kamu.uef.fi/en/tietopankki/guidance-and-support-in-doctoral-education/guidance-on-research-ethics/
UNESCO. (2004). Changing teaching practices: Using curriculum differentiation to respond to students' diversity. Retrieved April 22, 2024, from http://unesdoc.unesco.org/images/0013/001365/136583e.pdf.
UTSA. (2022, February). Tips for teaching students who are deaf or hard-of-hearing. University of Texas at San Antonio. Retrieved April 22, 2024, from https://www.utsa.edu/disability/faculty-staff/tips-teaching-deaf.html
van der Straaten, T. F. K., Rieffe, C., Soede, W., Netten, A. P., Dirks, E., Oudesluys-Murphy, A. M., Dekker, F. W., Böhringer, S., & Frijns, J. H. M. (2020). Quality of life of children with hearing loss in special and mainstream education: A longitudinal study. International Journal of Pediatric Otorhinolaryngology, 128, 109701. https://doi.org/10.1016/j.ijporl.2019.109701
Waltzman, S. B., Robbins, A. M., Green, J. E., & Cohen, N. L. (2003). Second oral language capabilities in children with cochlear implants. Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology, 24(5), 757–763. https://doi.org/10.1097/00129492-200309000-00012
Willberg, T., Buschermöhle, M., Sivonen, V., Aarnisalo, A. A., Löppönen, H., Kollmeier, B., & Dietz, A. (2016). The development and evaluation of the Finnish digit triplet test. Acta oto-laryngologica, 136(10), 1035–1040. https://doi.org/10.1080/00016489.2016.1175662
Winn, M. B., & Teece, K. H. (2021). Listening effort is not the same as speech intelligibility score. Trends in Hearing, 25, 233121652110276. https://doi.org/10.1177/23312165211027688

Wong, R. K. (2005). Advances in deafness management second language learning in Cochlear Implant Users October 9, 2005 Ripley K. Wong Speech Therapist in-charge queen mary. - ppt download. SlidePlayer. https://slideplayer.com/slide/3748110/

World Health Organization. (2018). Deafness and hearing loss. World Health Organization. Retrieved April 17, 2024, from https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. Ear & Hearing, 31(4), 480–490. https://doi.org/10.1097/aud.0b013e3181d4f251

Appendices

Appendix I. Ethical Consent Form

Date:_____/____/20__

ETHICAL CONSENT FORM

Investigator: Can Buldu
University of Eastern Finland

PURPOSE OF THE STUDY
The participant is invited to take part in a research study concerning the impact of listening effort on early language education of hard-of-hearing students. The participant's responses will be part of a data set which will further the understanding of listening effort measurements.

STUDY PROCEDURES
Should the participant consent, they will take part in a hearing test in which they will listen to sets of three digits read in Finnish, embedded in static noise. The participant will then repeat what they've heard. Simultaneously, they will complete a response time test in which they will click a button once they see a visual cue. The participant's name, age, occupation, hearing-related medical background, and level of Finnish will be recorded. The data will be terminated after the duration of the study. In any case of data reporting, the collected data will not be shared in any way that makes the identity of the participant known. The participant will remain anonymous, and in cases where a reference will be made, a pseudonym will be used to refer to the participant.

DURATION
The study will take one hour to complete.

VOLUNTARY PARTICIPATION
Please make note that participation is completely voluntary and dependent on the participant's wish and permission.

CONTACT INFORMATION
If you have any questions, please send an email to canbuldu@uef.fi.

CONSENT
"I voluntarily gave my consent to participate in this study. I have read and understood the information above."

Signature of participant

____

Appendix II. Interview Guide – Questions of Interest

1. What are the characteristics of the hard-of-hearing pupil's hearing device, school environment, preferences and familiarity with sign language and academic level?
2. What measures of support does the pupil receive?
3. What is the pupil's outlook towards his school life as part of mainstream education?
4. What is the pupil's outlook towards foreign language classes?
5. Is the pupil experiencing social isolation? If so, what measures are being taken?
6. How is the pupil's conduct during foreign language classes?
7. Are there any specific tasks or assignments that the pupil avoids?
8. How is the pupil's performance in tasks requiring concentration and multitasking?
9. Does the pupil experience increased fatigue from any classroom activities?
10. Is the teacher familiar with basic troubleshooting of the pupil's hearing devices?
11. Is the teacher employing any strategies to differentiate the content to accommodate the pupil's needs? If so, does the teacher have any suggested methods or classroom practices?
12. How is cooperation and communication conducted with the child's parents?
13. How is cooperation and communication conducted with medical and school support specialists?