LUND UNIVERSITY
Faculty of Medicine
Department of Speech, Phoniatrics and Audiology
Faculty of Medicine, Lund
Department of Audiology & Speech Pathology
University of Melbourne

Temporal resolution efficiency and auditory stream segmentation development

Marlin Johansson

Program of Audiology
Master thesis, 30 higher education credits
Supervisors: Gary Rance (PhD, University of Melbourne)
Jonas Brännström (PhD, Lund University)
Examiner: Ariane Laplante-Lévesque (PhD, Eriksholm Research centre & University of Linköping)

Abstract

Background: The ability to hear changes in a sound envelope is called temporal resolution and is assumed to have a key role in binaural speech segmentation. The development of temporal resolution is not fully explored, but studies indicate that it develops into adult-like efficiency at approximately 8-13 years of age, while binaural speech segmentation has an extended development up until approximately 13-16 years of age.

Method: A group of children (9-12 years of age) were compared to a group of 16 young adults. Temporal resolution was tested with a monaural non-speech amplitude modulation (AM) test at six modulation rates (10, 50, 75, 125, 150 and 250 Hz). Auditory stream segmentation was investigated with the binaural Listening in Spatialized Noise Sentence Test (LiSN–S). The purpose was to investigate in what age-span these abilities become adult-like, if there were development differences depending on modulation rate and if there was a correlation between the tests.

Results: A repeated measures ANOVA showed a significant difference between the two groups both on the AM-test (p=0.001) and LiSN-S test (p=0.002). A trend could also be seen that the 11-12 year old children (*n*=11) scored higher on the higher modulation rates compared to the lower ones. A positive significant correlation between LiSN-S spatial advantage score and the AM test at the modulation rates 50, 150 and 250 Hz was also found.

Conclusions: The results indicate that temporal resolution efficiency is still developing at 9-12 years of age, but that children can reach adult-like temporal resolution efficiency at about 11-12 years of age, especially at higher modulation rates (between 150-250 Hz). The last steps of the anatomical maturation is continuing until approximately the age of 11-12, which corresponds well to these findings. The correlation between the tests might indicate that an efficient temporal resolution is an advantage when use spatial cues.

Key words
Temporal resolution, auditory stream segmentation, Listening is Spatialized Noise-Sentence test, Temporal modulation transfer function, amplitude modulation, children, maturation, development

Abbreviations
AM = amplitude modulation
dB = decibel
IID = interaural intensity difference
ITD = interaural time difference
LiSN-S = Listening in Spatialized Noise–Sentence test
MLD = masking level difference test
SAM = sinusoidal amplitude modulated
SD = standard deviation
SNR = signal to noise ratio
SRT = speech reception threshold
TMTF = temporal modulation transfer function

TABLE OF CONTENTS

1. INTRODUCTION

Due to recent research efforts clinical audiologists have become increasingly aware of the problem with disturbances in the auditory processes. However, in order to understand these disturbances fully, the healthy developing auditory system also need to be studied. It is essential for the clinician who works with people with auditory deficits in whichever way, to know how auditory information travels from ear to cortex. To understand the deficits we must also have an understanding of how this system works in typical developing individuals.

Temporal resolution generally refers to the precision of a measurement with respect to time. In the context of the present study it is the ability of the auditory system to detect changes in the overall amplitude of a signal over time. Normal-hearing children are able to encode and use temporal envelope cues already at four or five years of age (Bertoncini, Serniclaes & Lorenzi, 2009; Hall & Grose, 1994). But the efficiency of using these cues have been found to develop further. Hall & Grose (1994) found the ability to develop until 8-9 years of age; Moore et al. (2011) found the skills to develop until 11-12 years of age and Banai et al. (2012) did not find temporal resolution to be adult-like in 11-13 year old children. Although high-quality studies within the audiology field, these studies do not give enough evidence to establish an age where normal development of temporal resolution become adult-like, because of different methods used and age groups evaluated.

Temporal resolution might also be an important key in the ability to segregate auditory stream information. Recent studies have found that there can be a correlation between scores on amplitude modulation (AM) detection tests (non-speech tests) and scores on the spatial 3D-test Listening in Spatialized Noise–Sentence test (LiSN-S, a speech based test) in children with neuropathy and autistic spectrum disorder (Rance, Ryan, Carew, Corben, Yiu, Tan & Delatycki, 2012; Rance, Saunders, Carew, Johansson & Tan, 2013).

Previous research indicate that children are not performing equal to adult results' until the age of 13-16 on the LiSN-S test (Brown, Cameron, Martin, Watson & Dillon, 2010). The indication from research that LiSN-S is developing into adult-like performance at a later stage than the AM test is not surprising, since it is a more complex task (with binaural speech stimuli).

1.1 Purpose

The present study investigates the development of temporal resolution with a monaural non-speech test and auditory stream segmentation with a binaural speech test in 9-12 year old children and compares the results to an adult group's result to investigate in what age-span these abilities become adult-like. This study can contribute to our knowledge of the healthy developing auditory system in children and is also a basis for further study on the disrupted auditory system.

1.2 Maturation of the auditory system

It is not fully understood which anatomical and physiological structures are responsible for temporal resolution and binaural speech segmentation. It is therefore meaningful to study the typical maturation of the auditory system and at what age the structures along the auditory pathway become adult-like. Cochlear function and anatomical development is essentially adult-like at or before the age of birth and the auditory brainstem pathway generates adult-like electrophysiological responses and structural development by the age of two (Ponton, Eggermont, Kwong & Don, 2000; Moore & Linthicum, 2007). Up to the midbrain and thalamus maturation proceeds rapidly and is completed by two or three years of age (Burkard, Don & Eggermont, 2007). However, the auditory cortex continues to mature into early childhood and maturation of the axons in the superficial layers of the cortex are developing up until the child is 11-12 years old (Burkard et al., 2007; Moore & Linthicum, 2007). Although all the anatomical structures are present at the age of 12, recent recordings from the temporal areas of the cortex indicate functional changes in both speech and tone processing in 10-18 year old children (Mahaja & McArthur, 2013). The adolescence is characterized by puberty and reorganization of brain functions subject to hormonal modulation of neural activity (Burkard et al., 2007). To summarize, previous studies indicate that anatomical maturation continues up until the child is 10-12 years old, but the functional processing of sound continue to develop across adolescence and into adulthood.

1.3 Temporal resolution

Being able to extract the temporal information from the acoustic signal is very important for speech processing and perception of music (Bellis, 2003). A spoken conversation is a time variant sequence of spectral shapes and accurate temporal processing and encoding is needed to understand what is being said (Bellis, 2003; Dreschler & Plomp, 1980). However, temporal processing does not only consist of temporal resolution, therefore an explanation of the expression is necessary. Temporal resolution refers to the detection of variations in the overall amplitude of the signal (Gelfand, 2009; Moore, 2008; Rance, McKay & Grayden, 2004; Viemeister & Plack, 1993). It should not be confused with the temporal fine structure i.e. the rapid pressure changes that carry the acoustic information. The acoustic wave consists of both detailed information (fine structure) and envelope information (what we refer to as temporal resolution). An illustration of the difference between temporal fine structure and the temporal resolution is presented in Figure 1.

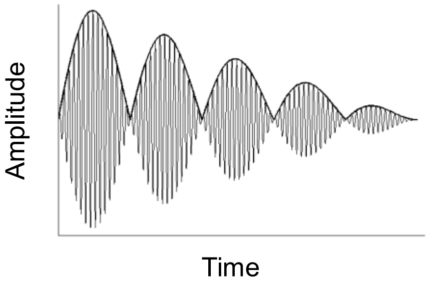

Figure 1. The difference between temporal fine structure and temporal resolution. The thick black lines following the envelope of the signals illustrate the overall envelope (temporal resolution) and the thinner rapidly changing wave-lines illustrate the fine structure.

1.3.1 How spectral cues influence measurements of temporal resolution

A fact to consider when measuring temporal resolution is that changes in the time pattern of a sound often are associated with changes in the sound's magnitude spectrum; e.g. the distribution of energy over frequencies or the spectral change. The shortest perceptible time interval is typically two to three milliseconds in young, normal developed adults when spectral cues are absent, and can be measured with detection of a brief gap between two sounds (gap detection) or amplitude fluctuations in a continuous sound (Gelfand, 2009; Rance et al., 2004). If spectral cues are present in the stimuli the subject can rely on the spectral cues and the results will not only reflect temporal resolution. When the signal contains spectral cues, a single click can be distinguished from a pair of clicks when the two clicks in a pair is only a few tens of a microsecond apart (Leshowitz, 1971; Moore, 2008). To avoid spectral cues a signal whose magnitude spectrum is not changed when the time pattern is altered have to be used i.e. a gap or a fluctuation in a white noise. Another approach is by adding extra background noise to mask the spectral changes. Taken together, when constructing a test to measure temporal resolution it is very important to avoid spectral cues, because when they are present, the results will not be accurate.

1.3.2 Temporal resolution measuring methods

Ronken (1970) measured temporal resolution by discrimination between two pairs of clicks with different amplitude (one high, one low) and different phase spectra but with identical long-term magnitude spectra. By varying the time interval between the clicks a threshold of 2-3 milliseconds for temporal resolution is received (Moore, 2008; Viemeister & Plack, 1993). Similar studies have been done with other stimulus durations and types of stimuli (Viemeister & Plack, 1993) most of them showing similar results to the values of the detection of a gap in broadband noise. Detection of how small gaps a subject can hear in a broadband noise (gap detection in noise) is probably the most frequently used technique to measure temporal resolution, because it provides a simple and convenient measure of temporal resolution (Moore, 2008; Viemeister & Plack, 1993). However, neither the discrimination between two click pairs nor gap-detection is enough to correctly reflect the advanced processing of envelope-temporal information in the auditory system (Moore 2008; Viemeister, 1979; Viemeister & Plack, 1993).

To only describe temporal resolution by a single number, the smallest detectable gap or the shortest duration of a click pair necessary for discrimination, is a limited approach. To understand the way temporal resolution works a model of the auditory system can be built instead. Most models of temporal resolution include band-pass filtering, reflecting the auditory filters in the cochlea. Each filter is followed by a nonlinear device, meant to reflect several processes in the peripheral auditory system, such as amplitude compression on the basilar membrane and neural transduction. The output is fed to another device that creates an approximating function to capture important patterns, a low-pass filter (Viemeister, 1979) or sliding temporal integrator (Moore, 2008; Plack & Moore, 1990). The last step is a decision mechanism which differs depending on task required, but its general purpose is to discriminate depending on approximating function from low pass filter/integrator. If it is a gap in the signal, the decision device is looking for a dip in the output, if a modulated sound is presented; the device tries to detect the AM of the sound (Moore, 2008; Viemeister & Plack, 1993).

A more general approach of measuring temporal resolution that better resembles the nonlinearity of the envelope-temporal processing in the auditory system is to measure the thresholds for detecting changes in amplitude as a function of modulation rate (Moore, 2008). Instead of generating a single number, this approach produces a model (or a function) that better resembles the complexity of the auditory system’s temporal resolution. The assembled result from AM-detection at different modulation rates is titled the temporal modulation transfer function (TMTF) and has the shape of a low pass filter. Thus, an accurate TMTF need several modulation rate thresholds. For low modulation rates the performance is limited by the level or amplitude resolution of the ear, rather than by temporal resolution. AM-thresholds are consequently independent of modulation rate up until 50 Hz. Above 50 Hz the ability to detect modulation is reduced with increasing modulation rate up to 800 Hz, thus the TMTF shows a low pass form (Viemeister & Plack, 1993). Above approximately 1000 Hz the modulation is hard to detect at all. The shape of the TMTF does not vary much depending on overall sound level, but the performance deteriorates at very low sound presentation levels (Viemeister & Plack, 1993). The capacity of detecting changes in AM is commonly calculated and presented as a function of different dB thresholds (Alcántara, Cope, Cope & Weisblatt, 2012; Eddins, 1993; Hall & Grose, 1994; Viemeister 1979; Viemeister & Plack, 1993). But the TMTF can also be presented as a single number. This is calculated with the 3 dB cut-off frequency; the frequency where the function has dropped 3 dB in amplitude from where it starts to fall. This cut-off frequency can be transformed into milliseconds with a mathematical transfer formula (Alcántara et al., 2012; Hall & Grose, 1994). Then the results can be compared to studies with less frequency specific measurements of temporal resolution (e.g. gap detection). Taken together the TMTF is a valid method of measuring temporal resolution because the results provide a better approximation of the auditory system function and it is also less level dependent compared to gap detection and other tests of temporal resolution.

1.3.3 Target stimuli

The AM-thresholds combined to build the TMTF are typically based upon sinusoidal AM noise (SAM-noise) and the results are presented as a negative dB threshold (Alcántara et al., 2012; Hall & Grose, 1994; Viemeister 1979; Viemeister & Plack, 1993). The modulation thresholds are usually expressed as 20 log (*m*), where *m* is the modulation index, or the amount modulation needed for detection (*m*=0 corresponds to no modulation and *m*=1 corresponds to 100% modulation). Figure 2 is a simplified example of a sinusoid (a curve of the equation y = sin *x*) and the relationship between the amount of modulation and the dB threshold. The stimulus is then built out of equal-amplitude sinusoids with random phase and compared to one or more noises without modulated sinusoids. Other approaches have been used in different studies, e.g. adding noise to one modulated sinusoid (Rance, 2010). However, using a broadband SAM-noise is probably the best known way to get rid of spectral and intensity cues, and is used to measure temporal resolution in the present study (Viemeister & Plack, 1993).

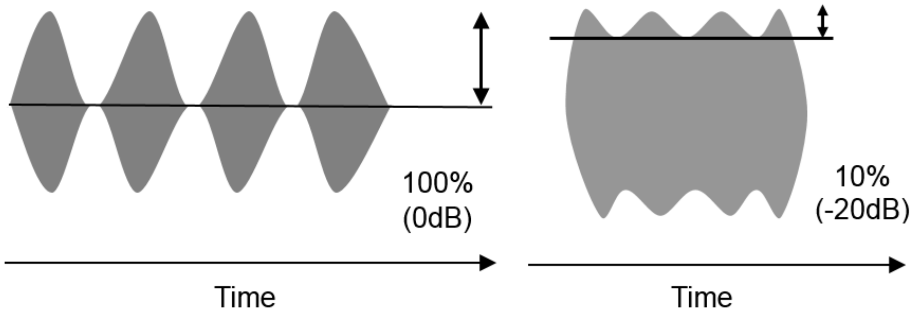

Figure 2. Two examples of the relationship between modulation rate and modulation threshold, measured with the formula 20 log (*m*) in dB, where *m* is the modulation index (10% corresponds to *m*=0.1).

1.3.4 The auditory system involving temporal resolution

As previously explained temporal resolution can be described with a model of four stages; band-pass filtering, a nonlinear device, a low-pass filter and finally a decision device. The first step is reflecting the auditory filters in the cochlea and the second step several processes in the peripheral auditory system, such as amplitude compression on the basilar membrane and neural transduction (Viemeister, 1979). All stages can be affecting temporal resolution, but according to Alcántara et al. (2012) there is a general consensus that the nonlinear device is most directly associated with sensitivity to modulation. Many children with autistic spectrum disorder have an impaired processing of AM and it is believed that this deficit is due to immature and/or delayed development of neural maturational processes (i.e. synaptogenesis, dendritic arborixation and myelination) (Alcántara et al., 2012), which leads to the conclusion that these processes needs to be fully developed in order to have a normal adult-like temporal resolution efficiency.

1.3.5 Past research of temporal resolution development

There are different ways of measuring temporal resolution. Hall & Grose (1994) measured TMTF with a SAM-noise in the same way as in the present study, but with a narrow band (200-1200 Hz wide) continuous noise. Hall & Grose (1994) examined the age-spans; 4-5, 6-7, 9-10 and 20-36 years of age. They tested 18 listeners; 4 in each child-group and 6 adults. The study indicated that sensitivity to modulation was reduced in younger listeners, at least up to 9-10 years of age. Hall & Grose (1994) could not find that the time constant and the 3 dB cut-off frequency of the TMTF’s varied significantly as a function of age. The authors suggest that even if temporal resolution may be adult-like at the age of 4, the level of efficiency of extracting cues to code temporal modulation improves with age up until the child is around 9-10 years old. From their results, however, it is hard to determine a specific age where temporal resolution is adult-like since the numbers of subjects in the age groups were small.

Moore et al. (2011) investigated AM-development in 6-11 year old children together with other auditory processing tests. They looked at groups of 6-7 (*n* =26), 8-9 (*n* =24), 10-11 (*n* =25) and young adults (no age specified in paper, *n* =21). A SAM-noise was used and the results were calculated as percentage modulation depth. They found similar patterns to Hall & Grose (1994); that the younger children needed larger modulation depths to detect the difference between stimuli than the 10-11 year olds and adults. They also claim that they found a delayed development of the 125 Hz-modulation rate relative to the 8 Hz-modulation rate. However, only two modulation rates were tested and as explained earlier; the 8 Hz-rate is limited by the level or amplitude resolution of the ear, rather than by temporal resolution. To summarize, the results from Moore et al. (2011) indicate that the sensitivity to AM still develops until the age of eight or nine, but because only two modulation rates were tested, the study lacks information about the general TMTF.

Banai et al. (2012) studied AM and frequency modulation (FM) in twelve children 8-11 years old, eleven 11-13 year old children, and adults 17-27 years old. The aim of the study was to investigate the developmental trajectories for the abilities to detect auditory amplitude and frequency modulation. The stimulus was a SAM-noise at modulation rates of 8, 64 and 125 Hz. The results indicate that mean performance for AM of wideband noise carriers was still not adult-like in the 11-13 year old group, even though the performance of the same group was adult-like on the FM-test.

Altogether results of previous TMTF studies indicate that temporal resolution efficiency still develops in eight to nine year old children. However, most recent research indicate that maturation of temporal resolution skills might even continue into teenage years.

1.4 Binaural hearing and binaural speech processing

Binaural speech processing is occasionally referred to as binaural interaction (Bellis, 2003). The general term is binaural hearing; the nature and effects of listening with two ears (Gelfand, 2009). Localization and lateralization of acoustic information, binaural release from masking, binaural fusion and detection of signals in noise are all functions that rely on binaural hearing (Bellis, 2003; Gelfand, 2009).

The ability to locate a sound in the auditory space is a complex task. The central auditory nervous system has to detect, perceive and compare small differences in arrival time and intensities between the two ears. If a sound is audible to the left, it will probably arrive to the left ear before the right ear. This small variance is called the interaural time difference (ITD). The head also acts as a barrier to high frequency sound waves, which also result in intensity differences between the two ears, which is called the interaural intensity difference (IID) (Bamiou, 2007; Cameron & Dillon, 2008).

1.4.1 Auditory stream segmentation

A sound rarely appears in isolation, we have sounds surrounding us all the time even though we do not notice them all. The sound sources may differ in terms of their location, intensity or their spectral or temporal complexity (Alain, 2007). It is very important for us to separate the important information from the sounds we do not want to hear. The ability to focus our attention on a particular sound object within an auditory scene presumes that the incoming composite acoustic wave has been ‘‘broken’’ or divided into its various elements (Alain, 2007; Bregman, 1990). Cameron & Dillon (2008) refer to the ability to segregate a target speech signal from simultaneously presented competing speech as auditory stream segmentation. It can also sometimes be referred to as selective attention (Shinn-Cunningham & Best, 2008).

The ability to understand speech in noise improves when the speech and noise sources are spatially separated (Bronkhorst & Plomp, 1988; Hirsh, 1950; Koehnke & Besing, 1996). This ability is based on binaural hearing and in particular ITD and IID. When the sound source of target stimulus and masking noise are the same, the interaural differences are similar. However, when the source locations are different, target stimuli and masking noise each have distinct interaural differences, which make the target stimuli easier to hear. Pickles (2008) refer to this ability as the “cocktail party effect” or spatial release from masking. These differences are believed to be the basis for improved speech intelligibility when the speech and noise sources are separated (Bronkhorst & Plomp, 1988; Koehnke & Besing, 1996; Pickles, 2008). However, the sound sources can also be distinguished from one another in terms of frequency information or spectral information, when two speakers are presented that have different formants or frequency spectrum in their voices for instance (Alain, 2007; Cameron & Dillon, 2008).

1.4.2 LiSN-S

The LiSN-S is used to evaluate the processes needed to understand speech in background noise, including binaural hearing (Cameron & Dillon, 2007). LiSN-S specifically measures the ability to segregate a target speech signal from simultaneously presented competing speech, previously referred to as auditory stream segmentation. The LiSN-S test was designed not only to determine if a child has an auditory stream deficit, but also weather the deficit is related to frequency related aspects of the speech signal or the physical location of the source in the auditory space (Cameron & Dillon, 2008). The classification of LiSN-S is a diotic test, because the same signal is presented simultaneously to each ear at the same time under earphones. When testing auditory stream segmentation with speakers in a free field the results can be limited by the subjects head movement which can affect the sound at the ear drum by several decibel (dB) (Cameron & Dillon, 2008). Furthermore the choice of loudspeaker, subject placement and reverberation in the test room are also factors that are better controlled when testing under earphones compared to in a sound field (Cameron & Dillon, 2008; Koehnke & Besing, 1996). LiSN-S presents a three-dimensional environment under earphones created by pre-synthesizing the speech stimuli with head-related transfer functions. Sentences are presented from 0° azimuth in competing speech in four different conditions, depending on location and vocal attribute of the competing sound source. When comparing different condition, advantage measures can also be calculated (Cameron & Dillon, 2007; Cameron & Dillon, 2008). How the competing sentences are presented in different conditions and the calculation of the advantage measures are depicted in Table 1. The result from each test condition is presented as a speech reception threshold (SRT) in dB.

Table 1. Introducing the different LiSN-S conditions and scores. Target sentences are always presented from 0° azimuth. Competing speech has the same vocal quality as the target voice, or different vocal quality.

LiSN-S conditions and advantage score	
DV90 =	competing speech from 90° azimuth, different vocal quality
SV90 =	competing speech from 90° azimuth, same vocal quality
DV0 =	competing speech from 0° azimuth, different vocal quality
SV0 =	competing speech from 0° azimuth, same vocal quality
Spatial advantage=	comparing the SV0 and DV0 to SV90 and DV90
Talker advantage=	comparing the SV0 and SV90 to DV0 and DV90

In absence of a segmentation deficit it is easier to hear the target stimuli if the sound sources differ in location or vocal quality because of a release from masking. Since the advantage measures calculate the difference between two test conditions performed by the same individual, the effects of linguistic ability between the subjects tested are minimalized. Cameron & Dillon (2008) tested a group of children with diagnosed learning or attention disorder (LD) and found no difference in performance between the LD group and a group without the diagnosis for any LiSN-S condition. This indicates that LiSN-S is not due to higher-order language, learning and communication functions. Taken together, LiSN-S provides a validated method of measuring binaural hearing that is similar to a situation children experience in their everyday environment. The different test conditions are also easy to compare to each other and compare between individuals.

1.4.3 LiSN-S compared to other tests of binaural hearing

McFarland & Cacace (2006) propose that a deficit in segregating auditory stream information should be measured with a dichotic test as one of many tests. Dichotic tests differ from diotic tests, such as LiSN-S, since different information is presented to the different ears simultaneously. However, both tests measure binaural hearing. An example of a dichotic test that is comparable to LiSN-S is the masking level difference test (MLD). Spatial release from masking is a feature in both tests and the results are calculated as a signal to noise (SNR) threshold in dB for both tests. Nevertheless MLD test is presenting the stimuli from one angle, with stimuli that is presented in different phase in each ear to cause the release from masking. Furthermore the stimuli are noise burst and tones, not speech stimuli. Cameron and Dillon (2008) compared the LiSN-S test with scores on other traditional auditory processing disorder (APD) assessment tools according to Bellis (2003); Dichotic Digits, 500 Hz MLD and Pitch Patterns Sequence test and found no correlation to any of these tests. The conclusion in the article was consequently that LiSN-S assesses unique auditory processes that are distinct from other APD assessment tools. It is believed that MLD test demonstrates good sensitivity to brainstem function, but that spatial cues such as ITD and IID continue to be processed in the auditory cortex and associated pathways (Cameron & Dillon, 2008).

1.4.4 Past research on LiSN-S

Cameron and Dillon (2007) tested LiSN-S on 82 children with normal hearing, 5-12 years of age. They found a trend of decreasing SRT thresholds and increased advantage scores as age increased across all LiSN-S performance measures. This indicates that segmentation skills develop beyond the age of eleven. Brown et al. (2010) investigated the North American Listening in Spatialized Noise – Sentence test (NA LiSN-S) in older individuals; one hundred and twenty normal hearing individuals between 12-30 years of age. They found a significant improvement in performance with age. However, the results of the different conditions developed into adult-like levels at different ages. There were no significant difference between the adults and children aged 13 and older on the SV0 condition. There were no significant differences between adults and children aged 14 and older on the talker and spatial advantage score. The DV90 condition developed into adult-like levels last, when the child was 16 years old.

In conclusion; there are not enough knowledge about the normal development of temporal resolution. Most studies on the subject has been using time efficient methods that are not frequency specific, thus not appropriately reflecting the complex processing of auditory information along the auditory pathway. Previous studies indicate that there might be a correlation between sensitivity to amplitude modulation and the auditory segmentation test LiSN-S in children with autism and neuropathy (Rance et al., 2012; Rance et al. 2013). However, this correlation has not been studied in normal developing children.

1.5 Research questions

In what age-span can we expect temporal resolution to be adult-like? Is the maturation for each modulation rate similar or can the results on some modulation rates be developed prior than others? Is there a connection between sensitivity to amplitude modulation and auditory stream segmentation in normal developing 9-12 year old children?

2. METHOD

2.1 Ethical clearance

This study was approved by the Research & Ethics Committee of the Royal Victorian Eye & Ear Hospital, Melbourne Australia. Written informed consents were obtained from all participants.

2.2 Subjects

Seventeen school aged children (six males) were recruited for this study. They were all native Australian speakers, lived in the Melbourne area and were children or friends of families of employees at the University of Melbourne. More girls than boys were tested, this was not intended; girls were easier assessable at the time of testing. Age at assessment ranged from 9 to 12 years (mean 10.7 years, standard deviation (SD) 0.7 years). Data were also collected from 16 young adults (six male), age at assessment 23-40 years (mean 28.8 years, SD 4.3 years). Adults <41 years of age were tested because research indicates that individuals over 60 years of age might have a decreased sensitivity to modulation detection (He, Mills, Ahlstrom, and Dubno, 2008). Eight of the young adults were native Australian speakers and lived in the Melbourne area and another two had English as a native language (originally from England and Scotland), but six of the adults did not have English as their first language (originally from France, Germany, Iran, Singapore and Sweden). When tested they all lived in the Melbourne area. Some were working and some were studying at a university level in Melbourne. No subject had any problem understanding written or spoken English and all included subjects had normal hearing, i.e. a mean threshold at 0.5, 1, 2 & 4 kHz of 20 dB HL or better on their best ear and 25 dB HL or better on their worse ear. Initially 17 adults were tested in the study but one of them had a sensorineural hearing impairment at higher frequencies in the left ear. Even a mild sensorineural hearing loss may influence once capacity of detecting AM and could also influence the LiSN-S results (Bacon and Viemeister, 1985; Glyde, Cameron, Dillon, Hickson and Seeto, 2013; Takahashi & Bacon, 1992). Hence he was excluded from the results, making the total adults 16 instead of 17. Ear examination with an otoscope presented unobstructed ear canals and normal tympanic membranes. The subjects were also asked for handedness and this information was used for deciding which ear to be tested in the AM-detection task, if left handed the left ear was tested and vice versa. This was conducted because the children group results' from the AM-test also provided another study with normal material. In the provided study this was the principle of choosing test-ear.

As LiSN-S required that each subject repeated a series of stimulus words, and as speech intelligibility impairments can affect LiSN-S, a phonetic profile was obtained for each subject using a version of the 'Articulation Survey' (Atkin & Fisher, 1996). Subjects were presented with several pictures, and were required to name the object on the picture. Each target phoneme in the word initial, medial and final position was elicited in the 'Articulation Survey Response Sheet'. Target phonemes (the 24 consonants used in Australian English) were recorded as either acquired or absent in each word position. Overall, the Articulation Survey revealed very few speech production errors. However, two children (sisters) did not pronounce one consonant correctly in one phoneme position. This was taken into account when scoring LiSN-S. If the child said the wrong consonant in a word, but it was obvious she misspoke because of the speech production error, she still received a correct score. Because LiSN-S is an adaptive method where the subject is introduced to 30 unique sentences for each test condition to estimate a SNR required for the listener to identify 50 % of the words in target sentence, the miscalculation arising from this is considered to be very small.

2.3 Procedure

Tests were typically performed in a quiet room during a single session. The background noise sound pressure level in the room was measured in connection to each session to make sure similar conditions was met, but the sound pressure level was not documented. The test battery is presented in Table 2. The only persons in the room when test was executed were the author (i.e. the examiner), the subject and occasionally a person who had already been a subject. The tests lasted approximately 35-50 minutes. A portable screening audiometer, Micromate 340 from Madsen Electronics and TDH 39 earphones were used to measure the subject's hearing. This equipment was calibrated according to AS 2586 and ISO 389-1 (ISO 389-1, 1998; Standards Australia, 1983) with a Bruel & Kjaer 4230 Sound Level Calibrator and a G.R.A.S. preamplifier 26 AC and a G.R.A.S. microphone 40 AN.

Table 2. The order and overview of the tests in the battery.

Test battery
1. Articulation survey
2. Screening with pure tones at 20 dB HL
3. AM-task
4. LiSN-S

The monaural AM-test was executed on a laptop as a computer game to find the minimum detectable depth for SAM-noise at modulation rates of 10, 50, 75, 125, 150 and 250 Hz. The modulation rates were presented from the lowest to the highest rate and the computers own soundcard was used. The earphones used were open HD650 earphones from Sennheiser. Because they are open earphones they give the subject a more natural sound and the examiner can hear the changes in amplitude and follow how well the subject performs. For each trial a 1000 ms segment was selected at random from generated noise; more information about the stimuli is found in Alcántara et al. (2012). In each trial, the subject listened to three consecutive noises and of these three two noises were identical and one noise was SAM-modulated. The stimuli were presented at a calibrated level of 70 dB SPL ± 1 dB SPL (clearly audible yet comfortable to listen to during the test session) and the silent interval between the noises in a trial was 500 ms. In a trial, after listening to the three noises, the subject were instructed to indicate the odd one out (i.e. which of the three noises they thought was different; the first, the second or the last or third). They were instructed to guess if they were uncertain about the correct response. Feedback was provided after each trial by a green light for correct and a red light for incorrect answer. The subject could then use the laptop and choose the box with the number that represented each noise themselves, or the examiner could click the box with the number that the subject told. However, the first interval on the 10 Hz tone was always performed with the examiner clicking the box and then the subject could choose to use the laptop if they understood the task. All subjects except one used the laptop themselves after the first trial. The modulation threshold was decided according to a three-down, one-up stepping rule to estimate the modulation depth necessary for 79.4 percentage correct detection (Alcántara et al., 2012;

Levitt, 1971). The step sizes and thresholds were based on the modulation depth in dB (defined in terms of 20log*m*, the relationship described in Figure 2). The step size of m variation was 5 dB at the start of each run, and then reduced to 2 dB after the first four reversals. Twelve reversals were obtained in a given run and the threshold estimate for that run correspond to the mean of m values of the eight last reversals. For each modulation rate subjects then received a threshold in dB and a SD for how much the subject's answers varied on each rate. If the SD level exceeded 3 dB the subject had to redo that specific modulation rate, according to Alcántara et al. (2012).

Binaural speech perception was assessed using the LiSN-S (Cameron & Dillon; 2009). Standard test earphones, HD215 from Sennheiser were used and connected to the headphone socket of a laptop via Phonak branded USB sound card, thus setting the sensitivity of the soundcard to a predetermined level to achieve the software's predesigned standard signal levels for LiSN-S (Cameron & Dillon; 2008; Cameron & Dillon; 2009). This alleviated the need of daily calibration (Cameron & Dillon, 2009). The test conditions was presented in the specific order recommended in Cameron & Dillon (2009); DV90, SV90, DV0 and finally SV0. Target sentences were presented at 62 dB SPL. Before each sentence a 55 dB SPL 1000 Hz 200 ms tone burst was presented to alert the subjects' attention. Competed stories for children were looped during playback to distract the subject from the target sentence. These stories were presented at a constant level of 55 dB SPL (combined level for two competing readers). The test was executed according to standard techniques by Cameron & Dillon (2009), however executed in a quiet, but not sound treated room. Up to 30 unique sentences for each test condition was presented to estimate a SNR required for the listener to identify 50 percent words correct for each condition. Each word in every sentence was scored separately by the examiner. The level of the target sentence was adjusted adaptively by the software to estimate the SRT according to Cameron & Dillon (2009). Two advantage measures were also calculated automatically by the software representing the dB benefit afforded by spatial or talker cues (c.f. Table 1) (Cameron & Dillon, 2008).

All subjects completed a hearing disability questionnaire at the end of each test session; abbreviated profile of hearing aid benefit (APHAB) (Cox, Alexander & Gray, 2003) and all children also completed a speech test with and without a FM-system. The data from the last tests are not reported here but some of the data is used in another study (Rance, Saunders, Carew, Johansson & Tan, 2013).

3. RESULTS

Independent samples t-test was used between parameters that could have affected the results (gender and handedness as grouping variable) for each of the different modulation rates, each LiSN-S conditions and the different LiSN-S advantage scores as test variables. No significant difference could be found on LiSN-S scores, which indicates that the parameters probably did not affect the result. On AM-test scores no differences could be seen for the rates and parameters, except that a significant result was found between handedness and the results on 250 Hz ($t(31)=2.6$, $p=0.01$). However, since only one left handed subject was tested, it is not enough evidence to say that it had an effect on the results.

3.1 The temporal modulation transfer function

Modulation detection thresholds from the AM-task were expressed in terms of 20log (*m*), were *m* is the modulation depth (*m*=1 corresponds to 100% modulation and *m*=0 to 0% modulation). Both the child and adult group showed TMTFs that are similar to the predicted low pass filter form. The data were curve fitted using a non-linear least-squares regression analysis formula in MATLAB R2012a according to Hall and Grouse (1994), who respectively based their model on Eddins (1993):

$f(x) = 10*\log10\ (1/(1+(a*x))\hat{}2)+c$

where modulation frequency is represented by x in Hz, *a* and *c* are estimated constants that could be derived from MATLAB in order to plot the fitted low-pass model in Figure 3. Because the performance is limited by the level or amplitude resolution of the ear for modulation rates lower than 50 Hz the result from the 10 Hz-rate was excluded from the curve fit in the figures.

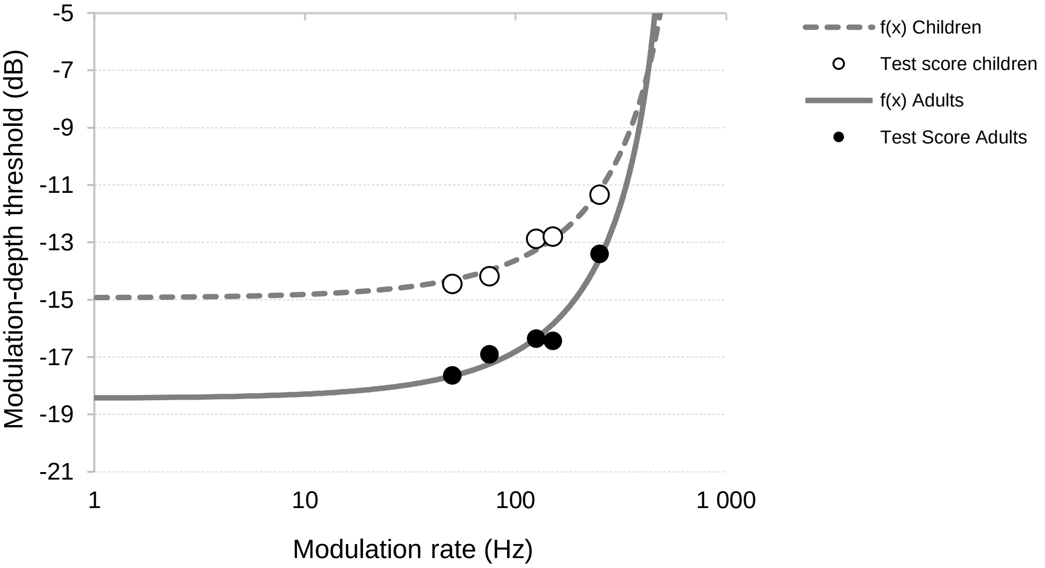

Figure 3. Mean group modulation-depth thresholds, expressed in decibels (dB), as a function of modulation rate (Hz) for the children (dotted) and adults (solid). The test results are curve fitted using a non-linear least-squares regression analysis formula. The x-axis is divided into a logarithmic scale, to be comparable to previous displays of TMTFs. The test scores on each modulation rate (white dots for children, black dots for adults) fit very well in with the lines f(x). This indicates that the results follow the pattern usually seen for temporal resolution testing (Alcántara et al., 2012; Eddins, 1993; Hall & Grose, 1994; Viemeister 1979; Viemeister & Plack, 1993).

The y intercept (when x=0 in the formula; $f(x) = 10*\log10\ (1/(1+(a*x))\hat{}2)+c)$) for the child group was -14.94 dB and -18.44 dB for the adult group, which indicates an increased sensitivity to modulation for the adult group compared to the children's group.

3.2 Amplitude modulation

Table 3 depicts the descriptive data for each modulation rate for each group. The SD was higher within the child group on every test rate compared to the adult group. This is also reflected in the min-max values, showing a larger range in results. When focusing on the mean results for each group, the mean modulation threshold was also higher for each frequency in the child group, indicating less sensitivity to modulation. The sensitivity for the groups is displayed in Figure 4.

Table 3. Descriptive statistics for different modulation rates divided into the two test groups. Independent-samples t-tests comparing the difference between the two groups of adults (N=16) and children (N=17), with Bonferroni correction are also displayed. *=significant at the 0.05 level, **=significant on the 0.01 level.

Modulation Rate (Hz)		**10**	**50**	**75**	**125**	**150**	**250**
Children (N=17)	**Mean (dB)**	-14.0	-14.5	-14.2	-12.9	-12.8	-11.3
	SD	2.2	3.1	4.3	4.0	3.6	3.1
	Min (dB)	-17.7	-19.1	-21.1	-21.3	-17.4	-17.2
	Max (dB)	-10.4	-6.8	-3.4	-5.1	-2.9	-5.6
Adults (N=16)	**Mean (dB)**	-16.8	-17.6	-16.9	-16.4	-16.4	-13.4
	SD	1.9	2.7	2.9	2.2	3.2	2.2
	Min (dB)	-19.4	-24.0	-21.8	-18.9	-20.1	-17.4
	Max (dB)	-12.6	-13.6	-11.6	-11.9	-8.7	-10.2
Difference between the groups	**T=**	3.9	3.2	2.1	3.1	3.1	2.2
	p=	0.003**	0.02*	0.3	0.03*	0.03*	0.2
	DF=	31	31	31	31	31	31

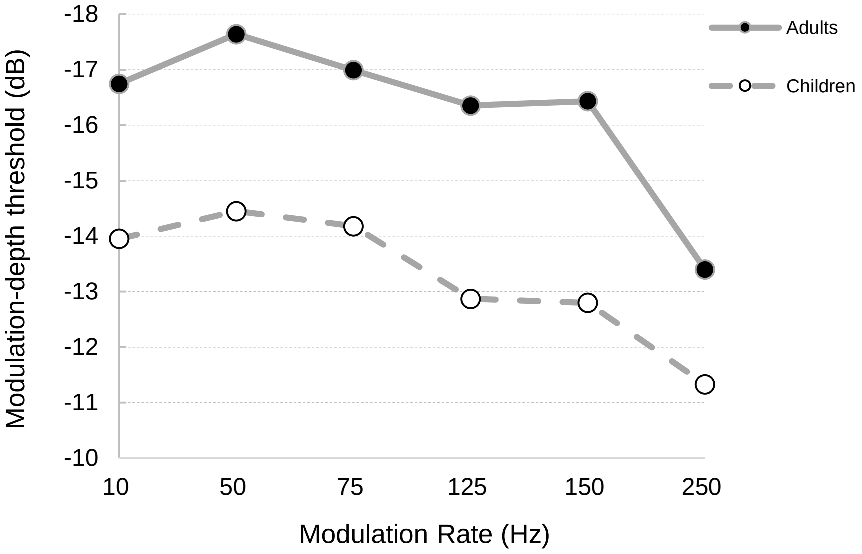

Figure 4. Mean group modulation-depth thresholds on each modulation rate. The adults (solid curve) displayed lower thresholds (i.e. better results) than the children (dashed curve) on each modulation rate (displayed by black dots for adults, white dots for children).

Repeated measures ANOVA was conducted on all test rates with group as between subject factor. The result displayed a significant difference between the groups (F=13.1, p=0.001). Independent-samples t-test was used to calculate the differences in results between the groups' mean results on each modulation rate. Because a series of tests were executed on the same individuals a Bonferroni correction was necessary to reduce the possibility of a type I error. The initial p-values were consequently multiplied by six. Table 3 shows the results of the t-tests, presenting significant differences between the groups on four of six modulation rates tested.

Pearson's correlation coefficients were calculated between age at assessment and each modulation rate threshold for all subjects. The analysis revealed a significant negative correlation between age and modulation rate on all test rates ($-0.4<r$, $p<0.05$), except at the 250 Hz modulation rate where it did to reach significance (see Figure 4). This indicates that with increased age the modulation threshold is lower (=increased sensitivity to modulation), except at the 250 Hz modulation rate. However, when correlating only the adults' age at assessment to the different modulation rate results for adults, as expected, no correlation between age and any modulation rate was found.

When excluding the younger children (9-11 years of age, *n* =6) from the result, only looking at the result of the children at 11-12 years of age (*n* =11) at all modulation rates, most of the 11-12 years olds are reaching what can be considered "adult-like" levels (presenting results within one adult SD from the mean adult value) at the higher modulation rates (Figure 5). The same pattern was not seen in either the 9-11 year old group or the adult group. The distribution of mean values for 50 Hz (the lowest rate tested for purely temporal resolution) and 250 Hz (the highest rate tested) is displayed in Figure 6, divided into the three age groups. Both Figure 5 and 6 indicate that the majority of the 11-12 year olds in the present study reaches what can be considered "adult-like" mean-values at the highest modulation rate tested. The distribution of the 11-12 year olds' results in the scatter plot in Figure 6 are also more similar to the adult groups' distribution of result on the 250 Hz rate, compared to the 50 Hz rate (the max and min results have approximately the same values for both groups).

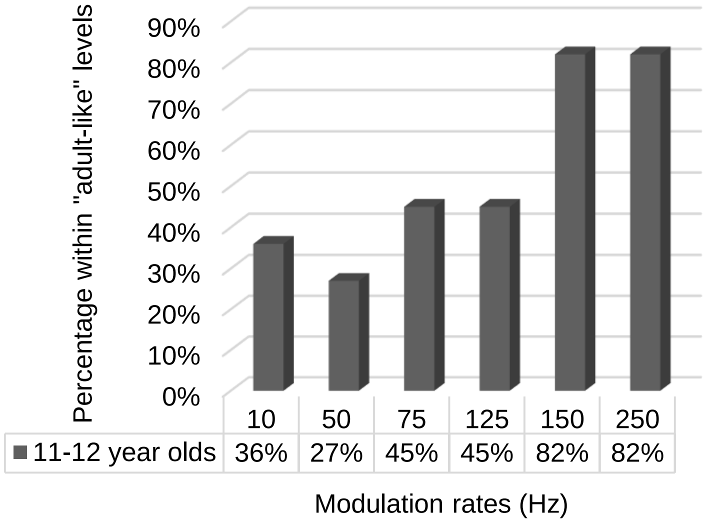

Figure 5. Percentage of 11-12 year old children (*n* =11) presenting modulation rate thresholds within one SD from the mean adult modulation rate threshold (what can be considered an "adult-like" result) at 10, 50, 75, 125, 150 & 250 Hz modulation rates. The figure show that the majority of 11-12 year old children performed at adult-like levels at the higher modulation rates.

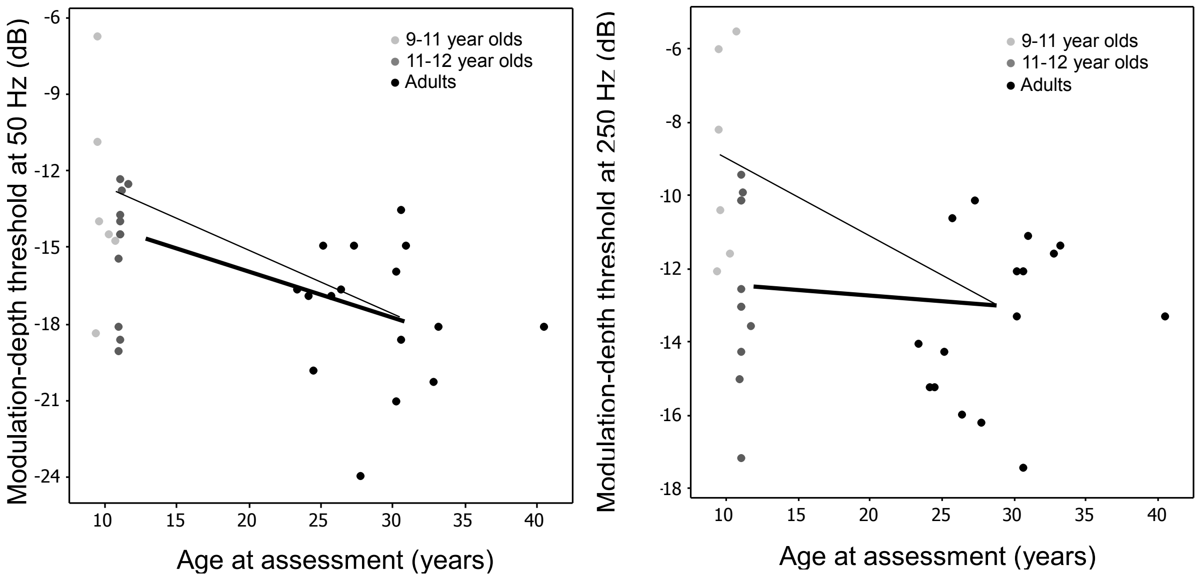

Figure 6. Distribution of mean results for all subjects divided into three groups based on age at assessment for the 50 Hz and the 250 Hz modulation rates. The thick black line is drawn between the mean for the 11-12 year old children (*n*=11) and the mean for the adults (*n*=16), the thin black line is drawn between the mean for the 9-11 year olds (*n*=6) and the mean for the adults (*n*=16). The figure indicates that the older children are reaching adult-like results at the 250 Hz rate to a greater degree than at the 50 Hz rate. The 250 Hz graph also indicate what can be considered to be a more adult-like distribution of results for the 11-12 year old children, compared to how the same group performed compared to the adults at the 50 Hz rate (the max and min results have approximately the same values for both groups at the 250 Hz rate).

3.3 Auditory stream segmentation

Table 4 and Figure 7 show the descriptive data for the different LiSN-S test conditions and the calculated advantage measures for each test group. In contrast to the AM-scores, the SD was slightly higher within the adult group than within the children's group on all LiSN-S test conditions. The adult group also showed a higher SD on talker advantage scores (but not spatial advantage scores) compared to the children. When viewing the means the adults showed lower thresholds on all listening conditions and higher advantage scores. A repeated measures ANOVA was also conducted on all LiSN-S test conditions and advantage measures with group as between subject factor. The result displayed a significant difference between the groups (F=11.1, p=0.002).

Table 4. Descriptive statics for different LiSN-S test conditions divided into the two test groups and t-tests comparing LiSN-S scores between the two groups of adults (*n*=16) and children (*n*=17), with Bonferroni correction are also displayed. *=significant at the 0.05 level, **=significant on the 0.01 level, ***=significant on the 0.001 level.

Test condition		**DV90**	**SV90**	**DV0**	**SV0**	**Talker advantage**	**Spatial advantage**
Children (N=17)	**Mean (dB)**	-13.6	-12.1	-4.9	-1.5	3.4	10.6
	SD	2.2	2.3	1.5	1.5	1.6	2.2
	Min (dB)	-16.7	-14.9	-7.4	-4.5	1.2	6.8
	Max (dB)	-8.0	-7.2	-0.8	1.9	6.5	13.5
Adults (N=16)	**Mean (dB)**	-16.6	-15.9	-8.9	-2.7	6.2	13.2
	SD	2.7	2.4	3.3	1.5	3.4	2.1
	Min (dB)	-19.6	-19.7	-15.0	-6.8	0.9	10.0
	Max (dB)	-10.2	-11.1	-3.4	-1.1	13.5	16.5
Difference between the groups	**T=**	3.5	4.6	4.4	2.3	-3.1	-3.5
	p=	0.009**	0.000***	0.001**	0.2	0.03*	0.01**
	DF=	31	31	21	31	31	31

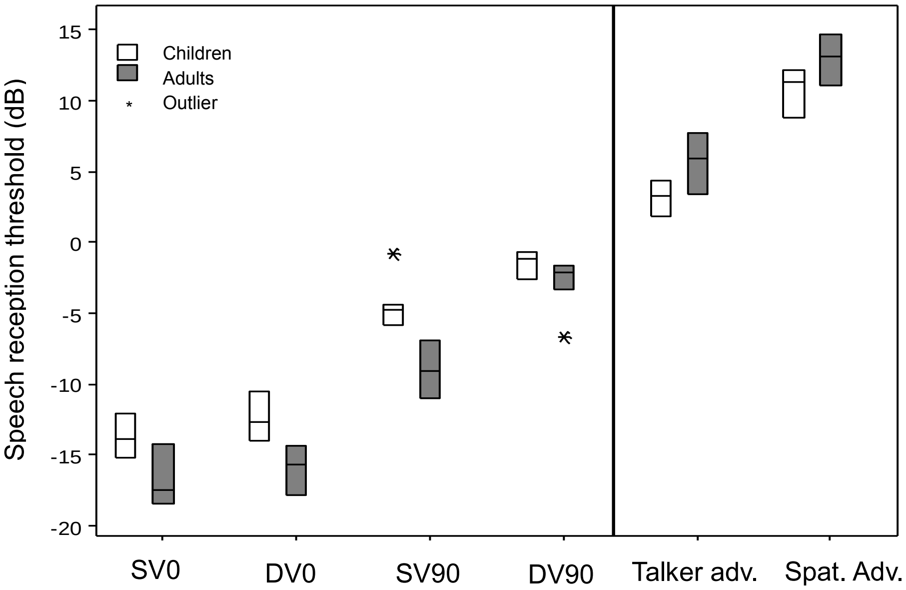

Figure 7. Speech reception thresholds for the four LiSN-S listening conditions and the two comparative measures; “Talker Advantage” and “Spatial Advantage”. The white bars represent the children’s group having higher thresholds (lower scores) on each test condition and lower thresholds (less benefit) on the advantage measures compared to the adult group with gray bars. The bars show median and quartiles.

Post-hoc independent samples t-test with Bonferroni correction presented significant differences between the groups on all test conditions, except SV0 where no talker or spatial cues were present. The advantage measures also displayed a significant difference between the groups. The t-tests comparing children’s results on LiSN-S to the young adults’ results can be seen in Table 4.

3.4 Correlation between auditory measures

A Pearson's correlation analysis was conducted on the LiSN-S scores and AM-test scores for both the child and the adult groups. No correlation could be seen between any of the LiSN-S scores and AM scores in the adult group. However, in the child group there was a correlation between some of the test conditions, displayed in Table 5. An example of the difference between the adult and child group correlations between the 250 Hz rate and spatial advantage scores is displayed in Figure 8. The child group showed significant positive correlations between the SV90 condition and the higher modulation rates tested (150 & 250 Hz), illustrating that a low SRT in the SV90 condition (i.e. better result) indicate a low modulation depth threshold at the higher modulation rates (i.e. better result). The children also displayed a negative correlation between spatial advantage and the 50, 150 and 250 Hz test rates, indicating that a high spatial advantage SRT (i.e. better result) is associated to a low modulation depth threshold at the 50, 150 and 250 Hz test rates (i.e. better result).

Table 5. Pearson's correlations between the child groups' LiSN-S and AM scores. Numbers in the table display the r value. *=significant at the 0.05 level, **=significant on the 0.01 level, ***=significant on the 0.001 level

Test condition	DV90	SV90	DV0	SV0	Talker advantage	Spatial advantage
AM 10 Hz	-0.3	-0.1	-0.3	-0.3	0.3	-0.1
AM 50 Hz	0.2	0.4	-0.3	-0.2	0.1	-0.5*
AM 75 Hz	0.2	0.2	-0.2	0.2	-0.3	-0.2
AM 125 Hz	0.4	0.4	0.2	0.2	-0.2	-0.4
AM 150 Hz	0.4	0.6**	0.01	0.2	-0.03	-0.6**
AM 250 Hz	0.1	0.6*	0.2	-0.2	-0.3	-0.7***

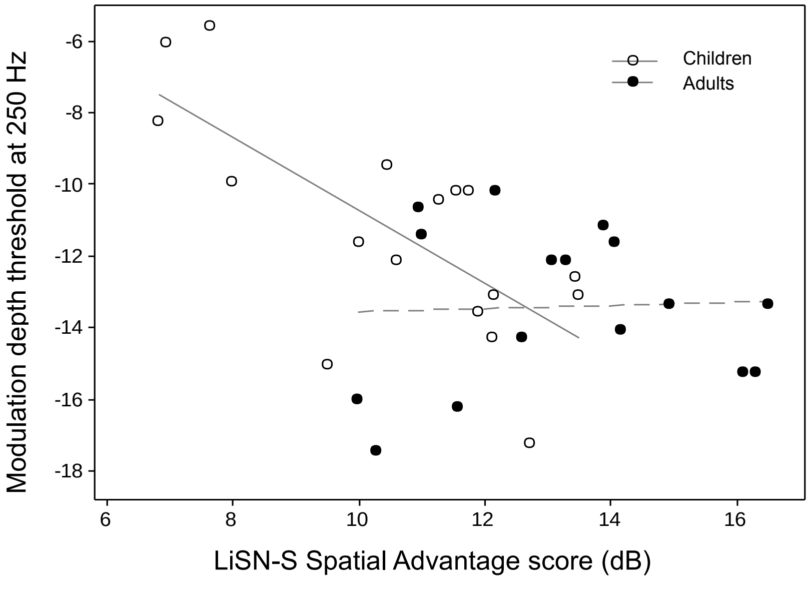

Figure 8. A significant relationship between modulation-depth threshold at 250 Hz and the LiSN-S spatial advantage score could be found on a Pearson correlation test (r=-0.5, p=0.002). However, when viewing the result within the groups, the children (solid curve) who had a lower modulation-depth threshold had a higher score on the LiSN-S spatial advantage task (r=-0.7, p=0.001) and no correlation could be seen within the adult group (dashed curve).

4. DISCUSSION

4.1 Method discussion

Because the present study was intended to result in a master thesis the opportunity of a big scale recruitment process of subjects was not possible, even if a larger sample size built on random choice could have made scientifically generalizations about the total population based on the results easier. The age-distribution within the child group of 9-12 year olds was not taken into consideration when constructing the tests, most children fell into the upper age span of 11-12 year olds (*n*=11), instead of the lower age span (*n*=6). The distribution of results within the child group was noticeably large on the AM-test, which a high SD in Table 3 indicates. The children with the poorest results on the AM-task generally belonged to the lower age span of 9-11 year olds (an example in Figure 6). Dividing the children into two or more groups, and also adding more children to each group, could have been beneficial when specifying in what age-span temporal resolution become adult-like.

The subjects were recruited by accidental sampling which could have influenced the results. The fact that all children tested were native Australian speakers and lived in the Melbourne area, while eight of the adults tested were native Australian speakers and another two had English as a native language, but six of the adults did not have English as their first language, could have influenced the results. However, the linguistic ability should not influence the results because the amplitude modulation task did not present speech stimuli and the advantage scores on LiSN-S is calculated as the difference between different test conditions for the same individual, thus excluding most of the effects of linguistic ability from the results. The SV0 is actually the only test condition that could have been affected by linguistic ability but it was the only LiSN-S test condition where no significant difference between the groups could be seen in the results, which probably indicates that no difference in linguistic ability between the groups was present. The social background in the adult group was diverse because people with different heritage was tested, some working and some studying in Melbourne. However, the education level was probably a little bit higher than seen for the average Australians, with three PhD-students tested. The adult group also consisted of people not living in Australia permanently, which might not be a proper representation of the Australian population.

Concentration levels could have affected the results, some children and adults were tested in evenings and some during daytime in the weekend, but the same procedure was used for all subjects. The presentation of the modulation rates in the AM-task was presented from the lowest rate to the highest rate and when the SD in the computer based program was above 3 dB the child or adult had to redo that specific test rate. This was based on the assumption that if you have a larger SD you might have lost your concentration level on that modulation rate. If the subject did not succeed to reach a SD below 3 dB in three test repetitions the best result was used. A randomization in the presentation of the modulation rates would have been a better approach to minimize any training and/or fatigue effects; that you become better at the task when practicing it a while and perform more poorly when becoming more tired. The reason why a training session was not offered was that limited time was scheduled for each subject, and LiSN-S and enough modulation rates to build a TMTF had a higher priority. Two children needed to redo the 75 Hz rate because of a high SD value but no adult had to redo that rate; this can be an indication of the shorter attention span in children which gives the adults and children different conditions, and why significance was not reached on these rates on some of the tests. This could have been avoided with a randomization of the presentation of modulation rates.

In a similar manner, differences in working memory capacity between the groups could have affected the performance. To include a working memory task would have proven that working memory did not have an effect on the result. On the other hand, the test battery was extensive at it was with six modulation rates tested and four LiSN-S subtests. The test battery might have had to be divided into two sessions if more tests were included, which would have complicated the testing procedure with the portable equipment.

The background noise sound pressure level in the room was measured in connection to each session both for children and adults, but not documented. To rule out the effect of noise level for each subject the documentation of this would have been useful to investigate the effect on the result. The TMTF does not vary much depending on overall sound level (Viemeister & Plack, 1993), thus should not be easily effected by noise level, but the amplitude resolution might have been affected at the 10 Hz rate tested, which the relatively high dB values in Figure 5 indicates (compared to the 50 Hz rate). The LiSN-S guidelines (Cameron & Dillon, 2009) recommend that the test is conducted in a sound proof environment, but as two groups in this study have been tested in the same way and we are interested in the difference, it should not have a large effect on the results.

4.2 Results discussion

Even though a limited number of modulation rates were tested in the present study, a typical low pass characteristic is displayed by the test rates in both curves in Figure 3; similar in form to TMTFs previously reported (Alcántara et al., 2012; Eddins, 1993; Hall & Grose, 1994; Viemeister 1979; Viemeister & Plack, 1993).

The y intercept for the child group in this study was -14.9 dB and -18.4 dB for the adult group, which indicates an increased sensitivity to modulation for the adult group compared to the children's group. This can be used to compare the overall sensitivity to modulation in this study compared to previous research. Hall and Grose (1994) expressed the dB thresholds in 20log ($1/m$) instead of 20log (m), and thus obtained a positive value for the modulation threshold. If transferred into 20log (m) the 9-10 year olds in Hall and Grose (1994) received a -16.9 dB y intercept and the adults a -17.8 dB y intercept. Compared to the present study the y intercepts for the adults were similar (-18.4 dB compared to -17.8 dB). However, the 9-12 year olds had an increased threshold in the present study compared to the 9-10 year olds (-14.9 dB compared to -16.9 dB). The difference can be due to a smaller number of subjects (4 children) in Hall & Grose (1994) or wider age-span in our group. It can also be due to a difference in test method, the difference in band width or the difference in test conditions. Because the y intercept is calculated from the point where the TMTF curve starts to increase in level (dB), the fact that we did not include modulation rates below 50 Hz into the TMTF can influence the result compared to previous studies. When the present study was compared to Alcántara et al. (2012), who used the same type of AM-test as in the present study, the y intercept was significantly higher; -25.7 dB for the 10-15 year olds tested. Most of the subjects are predicted to have adult-like processing at that age, but it is a high score even compared to the adult group in the present study, which also might be a reflection of the difficulty in comparing different studies due to dissimilar test conditions.

If comparing the specific modulation rates between studies, the mean 10 Hz modulation rate for the children in this study were -14.0 dB compared to approximately -17 dB for the 9-10 year olds in Hall and Grose (1994). The mean 10 Hz-threshold for the adults in the present study was -16.8 dB compared to approximately -19 dB in Eddins (1993) and -18 dB in Hall and Grose (1994). However, for the 50 Hz modulation rate (where only temporal resolution starts to have an effect) Hall and Grose (1994) found a -14 dB mean threshold for the 9-10 year olds tested, very similar to the -14.5 dB mean threshold found for the 9-12 year olds at 50 Hz in Figure 4. Moore et al. (2011) also used a SAM-noise and AM-test with a 125 Hz modulation rate. When calculating the results from percentage modulation depth into a dB mean threshold, the results at the 125 Hz rate was approximately -10.5 dB for 9-10 year old children and approximately -12 dB for the 11-12 year old children. This is a little bit higher, but not expressively different from the -12.9 dB mean threshold for the child group results in the present study (Figure 4). Taken together, the temporal resolution measured by the TMTF in the present study corresponds well to previous reports when comparing the modulation rates over 50 Hz. Why 10 Hz does not correspond to previous research can be due to the fact that we did not test in a sound proof condition (the 10 Hz rate is primarily based on amplitude resolution) or that we did not randomize the modulation rates or offered training before the test.

When performing repeated measures ANOVA across all test rates, a significant difference between the groups could be seen, with the adults hearing smaller differences in modulation than the children. These results indicate that there is a difference between the groups in the ability to detect amplitude modulation. This can be due to that temporal resolution efficiency is reduced in the younger group because of not fully matured auditory systems.

Maturation of the axons in the superficial layers of the cortex is developing up until the child is 11-12 years old (Burkard et al., 2007; Moore & Linthicum, 2007). Thus the last steps of the myelination process affecting the auditory pathway should be mature by the age of 12, and consequently the myelination development could be an explanation to why temporal resolution is not fully developed in the 9-12 year old group in general, but that some children reach adult-like levels at 11-12 years of age. Alcántara et al. (2012) believe that autistic children who display reduced temporal resolution efficiency could have a deficit in the natural myelination development, which also supports this hypothesis, because the autistic children tested had a similar TMTF pattern compared to controls, as the children compared to adults in the present study.

The child group's AM-results also point towards a trend that the higher modulation rates develop into adult-like results before the lower modulation rates. Moore et al. (2011) claimed that they found a delayed development of the 125 Hz-modulation rate, relative to the 8 Hz-modulation rate. However, only two modulation rates were tested and the 8 Hz-rate is limited by the level or amplitude resolution of the ear, rather than temporal resolution. The results in the present study indicate that the processing of temporal resolution on higher modulation rates might mature earlier than the lower modulation rates measuring temporal resolution and that some children actually show adult-like thresholds at 11-12 years of age (Figure 5 and 6). Because the auditory system is divided into frequency-specific channels up to the primary auditory cortex, it is not surprising that there might be a difference in the maturation of sensitivity to modulation depending on modulation rate.

The adult group also scored higher than the child group on LiSN-S, except at the S0-condition (Figure 7 and Table 4). The significant difference in spatial advantage and talker advantage probably indicate that the adult group benefits more from spatial and talker cues than the child group. The fact that no significant difference could be found in the S0-condition, when talker and spatial cues were absent, also consolidate this argument. The test without the cues actually tested the linguistic ability and the indication that there was no difference between the groups confirms the fact that we actually tested auditory segmentation and not linguistic ability. The present LiSN-S results also correspond well to previous findings of similar age groups, showing that 12 year olds did not have a fully developed ability to use spatial and talker cues (Cameron & Dillon, 2007; Brown et al., 2010). The spatial advantage mean for children at 9-12 years of age in Brown et al. (2010) was approximately 10 dB (somewhat lower for the 9 year olds and somewhat higher for the 12 year olds) which also corresponds well to the mean 10.6 dB for the child group in the present study. Furthermore the adults showed a threshold at 12 dB for spatial advantage in that study compared to 13.2 dB in the present study, which also could be considered a similar result. Talker advantage mean in Brown et al. (2010) was approximately 6 dB for the 9-12 years olds compared to 3.4 dB in the present study; this however is a poorer result showing that the children in the present study benefited less from talker cues. Talker advantage for the adults in the same study was approximately 9 dB compared to 6.2 dB in the present study. This difference can be due to the fact that the present study was not set in a sound proof booth. Talker cues are highly associated with the fundamental frequency. It is the fundamental frequency's average value, contour over time and temporal jitter that the listener use to separate two different voices (Rance et. al., 2012). Fundamental frequency is dependent upon amplitude and because it is the lowest frequency of a periodic waveform, it can be affected by low frequency sounds that are difficult to avoid in a quiet room without sound proofed walls.

Even though the anatomical maturation continues up until the child is 10-12 years old, the functional processing of sound continue to develop across adolescence and into adulthood. Because LiSN-S is a binaural speech test it is most definitely built upon more complex auditory connections than a monaural temporal resolution task. Thus it is not surprising that children cannot reach adult-like advantage scores on LiSN-S at the age of 12, because in order to reach adult-like auditory stream segmentation efficiency those complex connections need to be fully developed.

Comparable to Rance et al. (2012) the results in the present study demonstrated some correlations between LiSN-S scores and AM rate scores. The high AM detection rate (150 Hz) in Rance et al. (2012) was unrelated to SRT for the LiSN-S simulated-front listening conditions (DV0/SV0) and talker-advantage ($p>0.05$), comparable to the results in the present study where no correlation was found for those conditions either. Rance et al. (2012) did not find a correlation between the AM10 rate and any LiSN-S condition, also comparable to the results in the present study. However, Rance et al. (2012) found a correlation between all conditions requiring inter-aural location comparison (DV90, SV90 and spatial advantage) and the 150 Hz rate. The present study displayed significant correlations between the 150 Hz rate and two of the conditions; SV90 and spatial advantage. On the other hand the DV90 was not correlated to the 150 Hz rate. The DV90 was not correlated to the 250 Hz rate in the present study either, even though the SV90 was correlated to the 250 Hz rate. In comparison to the SV90 condition, cues to different talkers are also present in the DV90 condition which could be the reason why a correlation is not seen in the present study. Rance et al. tested 23 subjects, 9-55 years of age (27.4±14.1 years) with Friedreich's Ataxia and 12 subjects, 5-19 years of age (11.5±4.8 years) with Charcot–Marie–Tooth disease and compared the results to matched controls, which also is a different test group (and control group) compared to the present study. Rance et al. test groups' ability on both tests could be different from the children in the present study, because of the Friedreich's Ataxia and Charcot–Marie–Tooth disease and also the different age-spans tested. The difference between studies can also be due to different test conditions. Both studies were undertaken in a quiet room, however in Rance (2012) the background noise level was controlled and documented within <40 dB. Furthermore, the present study tested six modulation rates, compared to only two in Rance et al (2012), which could also have an effect on the outcome. However, despite the differences at the DV90 condition, the present study support the previous indication of a correlation between amplitude modulation detection and LiSN-S spatial advantage test in groups with reduced/not fully matured temporal resolution efficiency.

The correlation between the higher AM test rates and LiSN-S spatial scores can indicate that the same individuals just perform better or worse at both tests in the present study, maybe due to working memory capacity or general auditory processing performance. Another explanation could be that better temporal resolution efficiency leads to more benefit from spatial cues. This is not surprising because auditory stream segmentation due to location is actually a binaural processing task and therefore the auditory system is in need of good temporal information in order to use ITD properly. Because no correlation between auditory measures was documented in the adult group, this also indicates that if both amplitude modulation detection and the ability to use spatial cues are fully developed, there is no correlation between the tests. This would rather support the fact that both tests build upon some mutual ability than an effect of general performance or e.g. working memory.

4.3 Clinical implications

One of the purposes of this study was to contribute to our knowledge of the healthy developing auditory system in children. This study is one of the first studies to indicate that AM might develop into adult-like abilities at different pace depending on modulation rate, as previously suggested by Moore et al. (2011). However, more studies are needed in this area with an increased number of subjects.

This is also an important study for the design of classrooms and development of schools. With a larger number of pupils in a classroom, the noise level and competitive speech is increased, and as a consequence it is difficult for all children to listen to the teacher and classmates who try to speak their minds. With not fully matured auditory system and a hearing impairment it is very important that a child with a hearing impairment gets the help and support he/she needs for separating the desired signals from undesired noise. With advances in technology and especially communication systems that can reduce background noise and make it easier to focus on one speaker at the time, it can be of great value to investigate if all children could benefit from these systems, not only the hearing impaired. FM systems have been found useful for children with hearing losses since the beginning of the 1980s (Callingham, Kormandy & Weeks, 1983; Lewis, Feigin, Karasek, & Stelmachowicz, 1991), but lately also been found sucessful in treatment of auditory processing disorder (Sharma, Purdy, & Kelly, 2012) and other auditory perceptual disorders e.g. Friedreich's Ataxia (Rance, Corben, Du Bourg, King, & Delatycki, 2010). FM systems improve the SNR in noisy environments, so even if hearing impaired individuals benefit more from these systems, they help all individuals who wear FM systems to hear the signal better in noisy environments. Another more straightforward approach is to consider not putting too many pupils in one classroom and also make sure that noise, echoes, reverberation, and room modes do not interfere with the ability of the pupils to understand speech (Picard & Bradley, 2001).

The knowledge gained from this study is also important in the selection of hearing aids in the clinic and in the development of hearing aids for children. As a pediatric audiologist the focus is mainly based on audibility the first years of the hearing impaired child's life. Features to ease communication in background noise and binaural features to help compensate for the reduced ability of using binaural cues are often added later to the child's hearing aid. At what age advanced signal processing features are added in a hearing aid is different depending on the prescription of the selected hearing aid and/or when the audiologist chooses to introduce it to the child.

5. CONCLUSIONS

The present findings indicate that temporal resolution efficiency is not generally fully developed for individuals at 9-12 years of age. The findings also indicate that higher modulation rates might develop prior than other. The fact that the last steps of the anatomical maturation is continuing until approximately the age of 11-12 corresponds well to these findings. Furthermore, the children who could hear smaller changes in AM had a higher possibility to benefit better from spatial cues and score higher on the spatial processing tests (LiSN-S). However, no significant association was present in the adult group between any test conditions. This might indicate that more developed temporal resolution efficiency is a benefit in the ability to extract temporal cues and a stage towards adult-like auditory stream segmentation.

With advancements in technology and medical science the maturation of children's abilities to hear in noisy environment should be a highly prioritized subject within the audiology field. More studies are needed for older child groups (e.g. 12-14 year olds) and also an increased number of subjects, to confirm the specific age-span where temporal resolution becomes adult-like and to give more knowledge about the relationship between temporal resolution efficiency and spatial processing.

6. ACKNOWLEDGMENTS

I would like thank my supervisor Gary Rance who helped me design this study and made it possible for me to do the measurements and collect data. I would also like to acknowledge Elizabeth Barker who helped me with the subject recruitment and thank both her and Gary Rance for letting me do some of the testing in their home. I would like to thank Jonas Brännstöm my supervisor in Sweden for his support in the writing process. Also from Lund University I want to thank Anders Jönsson, for helping me establish the first contact with University of Melbourne and also helping me with the curve fit in MATLAB. Finally I would like to thank Peter Carew, Donella Chisari, Dani Tomlin and the rest of the staff at the Department of Audiology and Speech Pathology at University of Melbourne for giving me good advice and encouragement throughout the preparation of the study and in the testing process.

7. REFERENCES

Alain, C. (2007). Breaking the wave: effects of attention and learning on concurrent sound perception. *Hearing Research*, 229(1-2), 225-236. doi: 10.1016/j.heares.2007.01.007

Alcántara, J. I., Cope, T. E., Cope, W. & Weisblatt, E. J. (2012). Auditory temporal-envelope processing in high-functioning children with autism spectrum disorder. *Neuropsychologia*, 50(7), 1235-1251. doi: 10.1371/journal.pone.0044084

Atkin, N. & Fisher, J. (1996). *The Fisher Atkin articulation survey*. Parkville, Victoria: Royal Children's Hospital Speech Pathology Department.

Bacon, S. P. & N. F. Viemeister (1985). Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners. *Audiology*, 24(2), 117-134. doi: 10.3109/00206098509081545

Banai, K., Sabin, A. T. & Wright, B. A. (2011). Separable developmental trajectories for the abilities to detect auditory amplitude and frequency modulation. *Hearing Research*, 280(1-2), 219-227. doi: 10.1016/j.heares.2011.05.019

Bellis, T. J. (2003). *Assessment and management of central auditory processing disorders in the educational setting: From science to practice*. Clifton Park, NY: Thomson Learning.

Bregman A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Bronkhorst, A. W. & Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *Journal of the Acoustical Society of America*, 83(4), 1508-1518. doi: 10.1121/1.395906

Brown, D. K., Cameron, S., Martin, J. S., Watson, C. & Dillon, H. (2010). The North American Listening in Spatialized Noise-Sentences test (NA LiSN-S): Normative data and test-retest reliability studies for adolescents and young adults. *Journal of the American Academy of Audiology*, 21(10), 629-641. doi: 10.3766/jaaa.21.10.3

Burkard, R., Don, M. & Eggermont, J. J. (2007). *Auditory evoked potentials: Basic principles and clinical application*. Philadelphia, PA: Lippincott Williams & Wilkins.

Callingham, L., Kormandy, M. & Weeks, S. (1983). An FM Amplification system for conductive hearing loss (letter). *Medical Journal of Australia*, 2(11), 542.

Cameron, S. & Dillon, H. (2007). Development of the Listening in Spatialized Noise-Sentences Test (LiSN-S). *Ear and Hearing*, 28(2), 196-211. doi: 10.1097/AUD.0b013e318031267f

Cameron, S. & Dillon, H. (2008). The Listening in Spatialized Noise-Sentences test (LiSN-S): Comparison to the prototype LISN and results from children with either a suspected (central) auditory processing disorder or a confirmed language disorder. *Journal of the American Academy of Audiology*, 19(5), 377-391. doi: 10.3766/jaaa.19.5.2

Cameron, S. & Dillon, H. (2009) *Listening in Spatialized Noise - Sentence test (LiSN-S)* (Version 1.014). Murten, Switzerland: Phonak communications AG computer software.

Cameron, S., Glyde, H. & Dillon, H. (2011). Listening in Spatialized Noise-Sentences Test (LiSN-S): Normative and retest reliability data for adolescents and adults up to 60 years of age. *Journal of the American Academy of Audiology*, 22(10), 697-709. doi: 10.3766/jaaa.22.10.7

Cox, R. M., Alexander, G. C. & Gray G. A. (2003). Audiometric correlates of the unaided

APHAB. *Journal of the American Academy of Audiology*, 14(7), 361-371.

Dreschler, W. A. & Plomp, R. (1980). Relation between psychophysical data and speech perception for hearing-impaired listeners. *Journal of the Acoustical Society of America*, 68(6), 1608-1615. doi: 10.1121/1.392895

Gelfand, A.S. (2009). Essentials of Audiology. New York, NY: Thieme.

Glyde, H., Cameron, S., Dillon, H., Hickson, L. & Seeto M. (2013). The effects of hearing impairment and aging on spatial processing. *Ear and Hearing*, 34(1), 15-28. doi: 10.1097/AUD.0b013e3182617f94

Hall, J. W. & J. H. Grose (1994). Development of temporal resolution in children as measured by the temporal modulation transfer function. *Journal of the Acoustical Society of America*, 96(1), 150-154. doi: 10.1121/1.410474

He, N. J., Mills, J. H., Ahlstrom, J. B. & Dubno, J. R. (2008). Age-related differences in the temporal modulation transfer function with pure-tone carriers. *Journal of the Acoustical Society of America*, 124(6), 3841-3849. doi: 10.1121/1.2998779

Hirsh, I. J. (1950). The relation between localization and intelligibility. *Journal of the Acoustical Society of America*, 22(2), 196-200. doi: 10.1121/1.1906588

ISO 389-1 (1998). *Acoustics - Reference zero for the calibration of audiometric equipment. Part 1 - Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones*. Geneva: International Organization of Standardization 389-1.

Koehnke, J. & Besing J. M. (1996). A procedure for testing speech intelligibility in a virtual listening environment. *Ear and Hearing*, 17(3), 211-217. doi: 10.1097/00003446-199606000-00004

Leshowitz, B. (1971). Measurement of the two-click threshold. *Journal of the Acoustical Society of America*, 49(2), 462-466. doi: 10.1121/1.1912374

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49(2), 467-477. doi: 10.1121/1.1912375

Lewis, D. E., Feigin, J. A., Karasek, A. E. & Stelmachowicz, P. G. (1991). Evaluation and assessment of FM systems. *Ear and Hearing*, 12(4), 268-280.

Moore, B. C. J. (2008). Basic auditory processes involved in the analysis of speech sounds. *Biological Sciences*, 363(1493), 947-963. doi: 10.1098/rstb.2007.2152

Moore, D. R., Cowan, J.A., Riley, A., Edmondson-Jones, A. M & Ferguson, M. A. (2011). Development of auditory processing in 6- to 11-yr-old children. *Ear and Hearing*, 32(3), 269-285. doi: 10.1097/AUD.0b013e318201c468

Moore, J. K. & Linthicum, F. H. (2007). The human auditory system: A timeline of development. *International Journal of Audiology*, 46(9), 460-478. doi: 10.1080/14992020701383019

Picard, M. & Bradley, J. S. (2001). Revisiting speech interference in classrooms. *Audiology*, 40(5), 221-244. doi: 10.3109/00206090109073117

Ponton, C. W., Eggermont J. J., Kwong, B. & Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical Neurophysiology*, 111(2), 220-236. doi: 10.1016/S1388-2457(99)00236-9

Rance, G. (2010). Functional hearing deficits in listeners with auditory neuropathy spectrum disorder. *Revista de Logopedia, Foniatria y Audiologia*, 30(4), 204-210. doi: 10.1016/S0214-4603(10)70157-1

Rance, G., Corben, L. A., Du Bourg, E., King, A. & Delatycki, M. B. (2010). Successful treatment of auditory perceptual disorder in individuals with Friedreich's Ataxia. *Neuroscience*, 171(2), 552-555. doi: 10.1016/j.neuroscience.2010.09.013

Rance, G., McKay, C. & Grayden, D. (2004). Perceptual characterization of children with auditory neuropathy. *Ear and Hearing*, 25(1), 34-46. doi: 10.1097/01.AUD.0000111259.59690.B8

Rance, G., Ryan, M. M., Carew, P., Corben, L. A., Yiu, E., Tan, J. & Delatycki, M. B. (2012). Binaural speech processing in individuals with auditory neuropathy. *Neuroscience*, 13(226), 227-235. doi: 10.1016/j.neuroscience.2012.08.054

Rance G., Saunders K., Carew P., Johansson M. & Tan J. (2013). The Use of Listening Devices to Ameliorate Auditory Deficit in Children with Autism. *Journal of Pediatrics (corrected proof)*. doi: 10.1016/j.jpeds.2013.09.041.

Standards Australia (1983). *Australian Standard: Audiometers*. Canberra: AS2586-1983.

Ronken, D. A. (1970). Monaural detection of a phase difference between clicks. *Journal of the Acoustic Society of America*, 47(4), 1091-1099. doi: 10.1121/1.1912010

Sharma, M., Purdy, S. C. & Kelly, A. S. (2012). A randomized control trial of interventions in school-aged children with auditory processing disorders. *International Journal of Audiology*, 51(7), 506-518. doi: 10.3109/14992027.2012.670272

Shinn-Cunningham B. G. & Best V. (2008). Selective Attention in Normal and Impaired Hearing. *Trends in Amplification,* 12(4), 283-299. doi: 10.1177/1084713808325306

Takahashi, G. A. & Bacon, S. P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *Journal of Speech and Hearing Research*, 35(6), 1410-1421.

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *Journal of the Acoustical Society of America*, 66(5), 1364-1380. doi: 10.1121/1.383531

Viemeister, N. F. & Plack, C. J. (1993). *Time analysis: Human psychophysics*. New York, NY: Springer-Verlag.