MATURATION OF TEMPORAL PROCESSING IN CHILDREN: MEASUREMENTS USING SPEECH AND NON-SPEECH STIMULI

Saransh Jain, Bhavana Purigali Vasudevamurthy, Ashwini Pejathaya Raghavendra

Department of Audiology and Speech Language Pathology, Jagadguru Sri Shivarathreeshwara Institute of Speech and Hearing, Mysore, India

Corresponding author: Saransh Jain, Lecturer, Department of Audiology and Speech Language Pathology, JSS Institute of Speech and Hearing, Ooty Road, Mysore-25, Karnataka, India, e-mail: saranshavi@gmail.com

Abstract

Background: Auditory temporal processing is the ability of the nervous system to detect small variations in the duration of an acoustic stimuli. A substantial body of research is available on the development of various temporal skills, but temporal resolution abilities have not been well investigated in terms of speech and non-speech stimuli. The present study investigates the development of temporal resolution abilities in children.

Material and methods: A normative cross-sectional research design was adopted by administering a set of psychoacoustic tests involving both speech and non-speech stimuli. Six groups of 20 children each, aged 6–12 years, with a 1-year interval between each age group, were tested and the results were compared with those of 20 adults.

Results: The results revealed generally poorer performance of children on the entire test battery. Temporal modulation transfer function test scores, word recognition scores, and categorical perception of stop consonants matured by about 10–11 years of age. Gap detection test and time compressed speech test results showed maturation at around 8–9 years of age, whereas temporal change detection continued to mature even for the second decade of life.

Conclusions: Overall, maturation of temporal processing skills is reached by 10–11 years of age. This information is relevant when evaluating children with various processing disorders, and should also be considered when developing various assessment and rehabilitation protocols for children with special abilities.

Keywords: auditory processing disorders • psychophysics • psychoacoustics • auditory perception • speech perception • speech processing

---

DESARROLLO DEL PROCESAMIENTO DE LA ESTRUCTURA TEMPORAL DE LA SEÑAL EN LOS NIÑOS: MEDICIÓN CON LOS ESTÍMULOS VERBALES Y NO VERBALES

Resumen

Antecedentes: El procesamiento auditivo de la estructura temporal de la señal es la habilidad del sistema nervioso de reconocer pequeñas diferencias en la longitud de los estímulos acústicos. En actualidad se están llevando a cabo unos estudios detallados referentes al desarrollo de la habilidad del procesamiento de la estructura temporal de la señal, sin embargo, la posibilidad referente a la separación de la estructura temporal en relación con los estímulos del habla y otros estímulos aún no han sido examinados a fondo. Este trabajo estudia el proceso de desarrollo de las habilidades asociadas a la separación de la estructura temporal de los niños.

Material y métodos: Se ha aplicado el protocolo estándar del estudio transversal, que consiste en la aplicación de una serie de pruebas psicoacústicas, con utilización, entre otros, de estímulos de habla. En el estudio han participado seis grupos, cada uno de 20 niños, de 6 a 12 años, manteniendo un año de diferencia entre cada grupo de edad. Los resultados han sido comparados con los resultados de 20 adultos.

Resultados: Por lo general, los resultados de los niños han sido peores para toda la batería de pruebas. Los resultados de la prueba de modulación temporal, prueba de reconocimiento del habla y de la percepción categórica de vocales cerradas han demostrado que estas habilidades se desarrollan en la edad de 10-11 años. Los resultados de la prueba *Gaps in Noise* (detección de vacíos en el ruido) y para el discurso comprimido en el tiempo demuestran, que estas habilidades se desarrollan en la edad de 8–9 años; a su vez, la habilidad de detectar el cambio temporal se desarrolla incluso en la segunda década de la vida.

Conclusiones: En conclusión, la capacidad del procesamiento de la estructura temporal de la señal se desarrolla entre 10–11 años de edad. Esta conclusión se refiere a los niños que participaron en el estudio, con diversas disfunciones del procesamiento de sonidos y debería ser también tomada en cuenta al crear varios protocolos para la evaluación y rehabilitación para niños con habilidades especiales.

Palabras clave: trastornos del procesamiento auditivo • psicofísica • psicoacústica • percepción auditiva del habla • procesamiento del habla

ФОРМИРОВАНИЕ ОБРАБОТКИ ВРЕМЕННОЙ СТРУКТУРЫ СИГНАЛА У ДЕТЕЙ: ИЗМЕРЕНИЯ С ИСПОЛЬЗОВАНИЕМ ВЕРБАЛЬНЫХ И НЕВЕРБАЛЬНЫХ ИМПУЛЬСОВ

Изложение

Фон: Слуховая обработка временной структуры сигнала - это способность нервной системы распознавания небольших разниц в длине акустических импульсов. В настоящее время ведутся подробные исследования в области развития умения обработки временой структуры сигнала, однако возможности, касающиеся разрешения временной структуры по отношению к речевым и другим импульсам, не были еще тщательно исследованы. В настоящей работе исследуется процесс формирования умений, связанных с разрешением временной структуры у детей.

Материал и методы: Принят протокол нормативного профильного исследования, который заключается в применении комплекта психоакустических тестов, которые использовали, между прочими, речевые импульсы. В исследовании взяло участие шесть групп по 20 детей в возрасте от 6 до 12 лет с соблюдением одного года разницы между возрастными группами. Результаты были сравнены с результатами 20 взрослых людей.

Результаты: В общем результаты у детей были хуже для всего состава тестов. Результаты теста временной модуляции, теста на распознавание речи и категориального восприятия сплошных гласных сформировались приблизительно на 10–11 году жизни. Результаты тестов на выявление перерывов в шуме и на скомпрессированную речь во времени показали развитие умений приблизительно на 8–9 году, тогда как умение выявления временного изменения развивается во второй декаде жизни.

Итоги: Подводя итоги, умения обработки временной структуры сигнала развиваются к 10–11 году жизни. Этот итог относится к исследованным детям с разными дисфункциями обработки звуков и должен также учитываться при создавании разных протоколов оценки и реабилитации для детей со специальными умениями.

Ключевые слова: нарушения слуховой обработки • психофизика • психоакустика • слуховое восприятие • восприятие речи • обработка речи

KSZTAŁTOWANIE SIĘ PRZETWARZANIA STRUKTURY CZASOWEJ SYGNAŁU U DZIECI: POMIARY PRZY UŻYCIU BODŹCÓW WERBALNYCH I NIEWERBALNYCH

Streszczenie

Tło: Słuchowe przetwarzanie struktury czasowej sygnału to zdolność układu nerwowego do rozpoznawania niewielkich różnic w długości bodźców akustycznych. Obecnie prowadzone są szczegółowe badania w zakresie rozwoju umiejętności przetwarzania struktury czasowej sygnału, jednakże możliwości dotyczące rozdzielczości struktury czasowej w odniesieniu do bodźców mowy i innych nie zostały jeszcze dokładnie zbadane. Niniejsza praca bada proces kształtowania się umiejętności związanych z rozdzielczością struktury czasowej u dzieci.

Materiał i metody: Przyjęto protokół normatywnego badania przekrojowego, który polega na zastosowaniu zestawu testów psychoakustycznych, które wykorzystywały m.in. bodźce mowy. W badaniu udział wzięło sześć grup po 20 dzieci w wieku 6–12 lat, z zachowaniem jednego roku różnicy pomiędzy grupami wiekowymi. Wyniki były porównywane z wynikami 20 osób dorosłych.

Wyniki: Generalnie wyniki u dzieci były gorsze dla całej baterii testów. Wyniki testu modulacji czasowej, testu rozpoznawania mowy oraz kategorialnej percepcji samogłosek zwartych kształtowały się około 10–11 roku życia. Wyniki testów na wykrywanie

przerw w szumie oraz na mowę skompresowaną w czasie wskazały na rozwój badanych umiejętności około 8–9 roku życia, tymczasem umiejętność wykrywania zmiany czasowej rozwija się nawet w drugiej dekadzie życia.

Wnioski: Podsumowując, umiejętności przetwarzania struktury czasowej sygnału rozwijają się do 10-11 roku życia. Wniosek ten odnosi się do badanych dzieci z różnymi dysfunkcjami przetwarzania dźwięków i powinien być również uwzględniany przy tworzeniu różnych protokołów oceny i rehabilitacji dla dzieci ze specjalnymi umiejętnościami.

Słowa kluczowe: zaburzenia przetwarzania słuchowego • psychofizyka • psychoakustyka • percepcja słuchowa • percepcja mowy • przetwarzanie mowy

Background

Temporal processing is the ability of the auditory system to process time-related changes in a signal. It is an abstract term covering various underlying types of auditory perception. These processes can be broadly categorized as temporal discrimination, temporal masking, temporal pattern perception, and temporal resolution. Temporal discrimination is the ability to differentiate two or more acoustic signals in terms of duration; temporal masking involves perception of an acoustic signal in the presence of noise (where the noise can precede or follow the acoustic signal); temporal pattern perception is the ability of an individual to identify the pattern of presentation of an acoustic stimuli; and temporal resolution is defined as the ability to detect subtle changes in a sound signal with respect to time [1].

The auditory system has a restricted capability to follow a time-varying envelope. This limited ability of the auditory system to perceive envelope fluctuations is commonly known as temporal resolution [2]. Temporal changes in the envelope of speech signals provide various consonantal, supra-segmental, voicing, and speech boundary related information [3,4]. Adequate temporal resolution ability is necessary for speech perception because it provides a measure of perceiving vowels, consonants, syllables, and phrase boundaries. According to Schneider and Pichora-Fuller [5], temporal cues are important for speech intelligibility at least on two levels: segmental (phonemic) and supra-segmental (prosodic). At the segmental level, intrinsic temporal acoustic cues in speech sounds – like voice onset time, phoneme duration, burst duration, closure duration, formant transition duration, etc. – affect phoneme identification, whereas at the supra-segmental level, syllable rhythm and speed influence lexical and syntactic language processing.

Temporal resolution is a measure of both segmental and supra-segmental speech. At the segmental level, temporal resolution ability can be measured by using word recognition scores in the presence of noise [6]. Another way of assessing temporal resolution ability at the segmental level is through categorical perception of stop consonants. Elangovan and Stuart [7] found that the natural boundaries of between-channel gap-detection for speech stimuli may be measured using the categorical perception of voice onset time (VOT), as both share the same perceptual timing mechanism. Furthermore, the speed of lexical and syntactic language processing can be measured using the time compressed speech test [TCST; 8,9]. At the supra-segmental level, this ability can be measured using the temporal modulation transfer function [TMTF; 10],

temporal change detection (TCD), and gap detection threshold [GDT; 11].

Temporal resolution abilities are affected in individuals with hearing impairment, auditory processing disorders, and many neurological/developmental disorders. Ahmed et al. [12] found reduced temporal resolution abilities in children with specific language impairment. Researchers also reported poor temporal resolution of auditory signals in children with language impairment [13], dyslexia, and reading and writing disorders [14].

Temporal processing abilities are also influenced by auditory maturation. As with other auditory skills, temporal processing skills also show a developmental pattern. Maxon and Hochberg [15] investigated the auditory temporal integration abilities of normal hearing children aged 4–12 years and found a significant improvement in temporal integration thresholds as a function of age. Similarly, Davis and McCroskey [16] investigated the auditory fusion abilities in children within the age range 3–12 years and found an orderly increase in auditory fusion abilities from 3 to 8 years of age. Irwin et al. [17] investigated the auditory temporal acuity in 56 children of age 6–12 years using a gap in noise paradigm. They found temporal acuity improved with age until it reached adult-like values by 11 years. The effect of such maturation was also observed in other temporal processing skills assessed by tests like forward and backward masking [18], co-modulation masking release [19], and voice onset time [20].

Immature temporal processing abilities may in part explain the limited capacity of children to identify small variations in speech [21]. However, there is a discrepancy in the literature regarding the minimum age of maturation for various temporal skills. The maturation age for auditory temporal processing skills ranges from as low as 7 years to as much as 12 years depending on the research study. Further, although this maturation data was derived from specific tests, no systematic study considering all aspects of temporal resolution ability is available. Another limitation of the available literature is the lack of adequate norms. In all studies, a control group was considered and compared with an experimental group, as no specific norms were available. Lastly, most previous studies have considered the development of temporal resolution abilities using tonal or noise stimuli; however, there is a dearth of literature considering speech stimuli. Since the ultimate motive in assessing temporal resolution ability is to measure an individual's speech perception abilities, it is important to assess temporal processing abilities using speech stimuli. Thus, there is a need to investigate the development of temporal resolution abilities in children by

Table 1. The group wise distribution of the participants with respect to age

Group	Age range	Mean age	Number of participants
Group 1	6–7 years	6.4 years	20 (10 M + 10 F)
Group 2	7–8 years	7.6 years	20 (10 M + 10 F)
Group 3	8–9 years	8.6 years	20 (10 M + 10 F)
Group 4	9–10 years	9.4 years	20 (10 M + 10 F)
Group 5	10–11 years	10.6 years	20 (10 M + 10 F)
Group 6	11–12 years	11.6 years	20 (10 M + 10 F)
Group 7	18–25 years	23.5 years	20 (10 M + 10 F)

considering a set of tests involving both speech and non-speech stimuli. It was hypothesized that temporal resolution abilities should show a systematic development with age over the early years of life. Hence, in a quest to establish normative data, the present study aimed to investigate the development of various temporal resolution abilities for speech and non-speech stimuli in normally developing children over the age range 6–12 years.

Material and Methods

Subjects

A cross-sectional normative research design was employed by considering, along with 20 adults, a cohort of 120 typically developing children within the age range 6–12 years. There were six age groups of 20 children each, with 1-year interval between each group. Thus, six pediatric and one adult group were studied. The children in each age group were equally divided in terms of gender and efforts were taken to select children which were normally distributed in that particular age range (Table 1).

A formal audiometric screening was carried out on each participant to rule out the possibility of hearing loss affecting the test scores. Only participants who passed hearing screening (PTA ≤15 dB; SRT ±10 dB of PTA; SIS ≥90%) [22] were considered. Formal speech and language screening by a qualified speech language pathologist was also performed and participants with inadequate/delayed/deviant speech and language abilities were excluded from the study. All the participants were native Kannada speakers and were also screened for neurological or psychological impairment by a qualified school psychologist, as a part of health screening in affiliated schools from our organization. An informed written consent was obtained from either the participants (in case of adult participants) or from parents/teachers of the participants (on behalf of pediatric participants). The purpose of the study was also briefly explained before commencement of the study. Prior approval from the institutional review board to study human subjects was obtained.

Tests and stimuli

A series of tests including TMTF, GDT, TCD, WRS in continuous and interrupted noise, categorical perception (CP) of stop consonants, and TCST were administered. TMTF, GDT, and TCD were assessed using non-speech

stimuli, whereas WRS, CP, and TCST were assessed using speech stimuli.

Non-speech tests and stimuli

The TMTF, TCD, and GDT were assessed by following the maximum likelihood procedure (MLP, a Matlab toolbox for psychoacoustic experiments) [23]. The TMTF was measured using 8, 20, 60, and 200 Hz modulation depth. The stimulus was 500 ms Gaussian noise sinusoidally modulated at a specific modulation depth with 10 ms raised cosine ramps at onset and offset. The modulation depth was expressed logarithmically (dB). The participant's task was to identify the interval containing the modulated noise in a 3AFC paradigm.

TCD was measured using a sequence of five tones of 100 ms each separated by a temporal gap. All the tones were similar in frequency, with the first three tones gated with 10 ms onset and offset cosine ramps. The participant's task was to identify the interval with the irregular rhythm in a 3AFC paradigm.

The GDT was measured using a 750 ms Gaussian noise having a 0.5 ms cosine ramp at the beginning and the end of the gap. The participant's task was to identify the interval containing the gap in a 3AFC paradigm. The complete testing, involving TMTF at four modulation depths, TCD, and GDT, was carried out using 30 trials per block for a total of 5 blocks.

Speech tests and stimuli

WRS was administered in the presence of continuous and interrupted noise at various SNR levels. Initially a total of 450 words were selected from a kindergarten vocabulary and were given to 10 native Kannada speakers to rate the words on a 5-point familiarity rating scale. All raters were pre-primary school teachers who were familiar with kindergarten vocabulary. After the rating, only those words with a score of 4 or 5 (familiar or very familiar) were selected for testing. Thus, a list of 200 Kannada bisyllabic words present in kindergarten vocabulary were selected and divided into 10 lists of 20 words each. Five of these lists were associated with continuous noise and the remaining five were associated with interrupted noise, using COLEA software (Matlab software tool for speech analysis) [24] with each list at an SNR of –9, –6, –3, 0, and +3 dB. The participant's task was to identify the correct words in the

presence of noise, and the minimum SNR level at which the participants could identify 50% of the words correctly (SNR50) was considered as the threshold. The word lists were presented randomly and none of the word list was presented twice to the same participant, in order to reduce familiarity effects.

The stimulus for the categorical perception was a |ba|–|pa| sequence. The voice onset time (VOT) of |ba| was systematically varied on a 13-point scale in such a way that it was perceived as |ba| at one end-point and |pa| at the other. The VOT was defined as the onset of the voicing to the onset of the formant transition without affecting the burst. The stimulus was a natural speech token recorded by a female speaker at a sampling rate of 44100 Hz. The VOT was truncated using Praat software (version 5.1.41) [25] in 5 ms steps over a range of 0 to 60 ms. Each time the VOT was truncated, a silent gap of the same duration was inserted to simulate bilabial stop closure. The remaining acoustic and temporal parameters were kept constant across the continuum. To increase reliability, each stimulus along the continuum was presented 5 times in random order to each participant. The participant's task was to identify the stimulus in a single interval forced choice paradigm and the categorical boundary was marked. This boundary was defined as the shift in the perception from voiced |ba| to voiceless |pa|.

TCST was also administered using three- to four-word sentences present in kindergarten speech. The familiarity of the sentences was also done by employing a similar procedure to that used for the WRS familiarity check. A final list of 50 sentences was selected and divided into five lists of 10 sentences each. One list was presented with no temporal compression, while the other four were presented with 50%, 60%, 70%, and 80% time compression. The stimulus was compressed using Audacity software (version 1.3.14-beta) [26] and the maximum compression at which the participants were able to identify 50% of the sentences correctly was considered the threshold. The sentence lists were presented in random order and none of the lists was presented twice to any participant.

Procedure

The complete testing was carried out on a Dell Inspiron i7 personal computer equipped with Sennheiser HD220 circumaural headphones. The output of the headphones was measured and controlled using a B&K 2268 (mediator) sound level meter to the most comfortable level of each participant. The complete testing was carried out in a separate room in the school (for group 1 to group 6) which had minimal environmental noise and reverberation. Because the most comfortable level was measured in the sound-treated audiometric setup, whereas testing was carried out in a quiet but not sound-treated room, a correction factor of 10 dB was added to the most comfortable level. The correction factor was obtained by measuring the average most comfortable level difference on a set of adult normal hearing individuals in an audiometric situation and in a general quiet environmental situation (this information was originally collected for another research study, but was utilized in this study as well). This level was also correlated with the subjective perception of loudness

during the practice trial, and wherever necessary the output was increased or decreased slightly to match the participant's comfort. The entire testing for the non-speech stimuli was performed in a single sitting, with a break of 5 minutes after each test, to reduce habituation and prevent auditory fatigue. The remaining testing using the speech stimuli was completed on the following day, as the time required for testing was too long and tiring, especially for the children, to carry out in a single day.
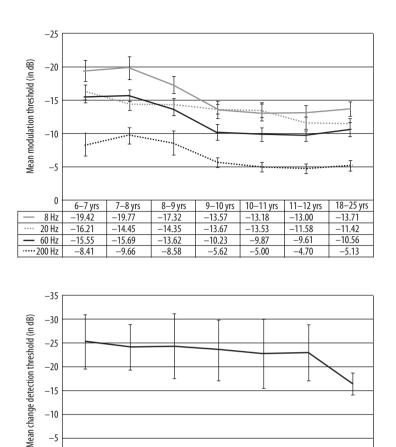
Data analysis

The results obtained from each test were subjected to appropriate descriptive and inferential statistics. Multiple group-wise comparison was carried out using one-way ANOVA and the data was subject to Bonferroni's post hoc test to establish the developmental trajectory. However, some specific statistical procedures were employed to analyze the results. In the word recognition test, in order to see the interaction effect of age, type of noise, and SNR, 3-factor multiple ANOVA was carried out. Similarly, in the categorical perception test, the categorical boundary was obtained by estimating the 50% point on the psychometric curve using a logistic regression model with either linear or nonlinear interpolation.

Results

Temporal modulation transfer function

The data obtained from the tests were compared across age groups, using one-way ANOVA. The TMTF was measured using 8, 20, 60, and 200 Hz modulation depths. The results revealed the poorest mean TMTF scores (for all modulation depths) for the 6–7 years age group and better mean scores for the 10–11 years age group (Figure 1). After 11 years the mean scores were approximately the same as with adults. The statistical significance of these differences was estimated by looking at the effect of group on responses. A significant between-group effect was observed at modulation depths of 8 Hz [$F(5,119)=47.70$; $p<0.05$], 20 Hz [$F(5,119)=32.47$; $p<0.05$], 60 Hz [$F(5,119)=46.76$; $p<0.05$], and 200 Hz [$F(5,119)=13.79$; $p<0.05$].

Bonferroni post hoc analysis also revealed significant differences in TMTF scores (for all modulation depths) between the youngest and oldest paediatric group. However, there was no significant difference from 6 years to 9 years of age, even though the mean scores systematically improved, indicating a developmental trend. Similarly, from 10 to 12 years of age, no significant difference was observed, although a statistically significant difference was observed between 9 to 10 years for modulation depths of 8 Hz and 20 Hz. For a modulation depth of 60 Hz, this difference was only observed between 9 to 11 years but not between 10 to 11 years. On the other hand, no such significant group differences were observed between 9 years to 11 years for a modulation depth of 200 Hz; nevertheless, at this modulation depth a significant difference was observed from 7 years to 11 years of age. These results indicate systematic improvement in TMTF scores up to 10 years, after which it reaches adult values.

Figure 1. Mean temporal modulation transfer function thresholds at different modulation depths (dB, vertical axis) and ages (horizontal axis). Modulation frequency is parameter (black/grey lines). More negative values indicate better performance. Bars indicate standard deviations

	6–7 yrs	7–8 yrs	8–9 yrs	9–10 yrs	10–11 yrs	11–12 yrs	18–25 yrs
— 8 Hz	−19.42	−19.77	−17.32	−13.57	−13.18	−13.00	−13.71
⋯ 20 Hz	−16.21	−14.45	−14.35	−13.67	−13.53	−11.58	−11.42
— 60 Hz	−15.55	−15.69	−13.62	−10.23	−9.87	−9.61	−10.56
⋯ 200 Hz	−8.41	−9.66	−8.58	−5.62	−5.00	−4.70	−5.13

Figure 2. Mean temporal change detection thresholds (dB, vertical axis) in different age groups (horizontal axis). More negative values indicate better performance

	6–7 yrs	7–8 yrs	8–9 yrs	9–10 yrs	10–11 yrs	11–12 yrs	18–25 yrs
— CDT	−25.43	−24.37	−24.52	−23.81	−22.98	−23.17	−16.52

Temporal change detection

The results of TCD were different from that of TMTF scores. In this particular test, adult-like values were not achieved even by 12 years of age. The mean TCD thresholds for 6–7 year old children were far poorer than those of adults. As seen in Figure 2, only small changes in mean thresholds were observed with increasing age, although these changes had high standard deviations. A between-group comparison revealed no significant difference in detection threshold from 6 years to 12 years of age [$F(5,119)=3.74$; $p>0.05$]. However, pair-wise comparison between the paediatric and adult population revealed a significant difference, with the change detection thresholds of adults found to be far better than those of children. The multiple pair-wise comparisons using the Bonferroni post hoc test also revealed no significant difference from 6–7 years of age to 11–12 years. The results also revealed a significant difference between 12 years and adulthood, indicating that the scores had not reached mature values by this age.
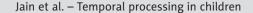
Gap detection test

A similar trend to TMTF was observed in the GDT results, where it was found that the mean scores of the youngest age group were the poorest and the scores reached adult values by 8–9 years of age. One-way ANOVA revealed a significant main effect between groups [$F(5,119)=8.69$; $p<0.05$]. Bonferroni post hoc comparison also revealed a significant difference from 6 years to 12 years of age. However, contrary to the TMTF scores, the GDT scores reached adult-like values as early as 8 years of age, after which only small improvements were observed.

Word recognition scores

Further analysis of temporal resolution using speech stimuli was carried out by administering WRS, CP, and TCST. The mean WRS scores for SNR50 (50% correct responses) [27] in the presence of continuous noise were obtained at −3dB SNR for the 6–7 years age group, whereas it reached scores comparable to those of adults by 10–11 years of age. At this age group, SNR50 was obtained at −9 dB SNR and remained steady with further increase in age.

The inverted bars in Figure 4 represent WRS scores below SNR50. The inverted bars are evident for poor SNR (−9 and −6 dB), and scores greater than 50% correct are only obtained at 10–11 years of age. At good SNR levels, 50% correct word identification is achieved even by 6–7 years of age. A somewhat similar trend is observed in Figure 5 for interrupted noise, but the age at which 50% correct word identification was reached occurred as early as 8–9 years of age, even for poor SNR levels (−9 dB).
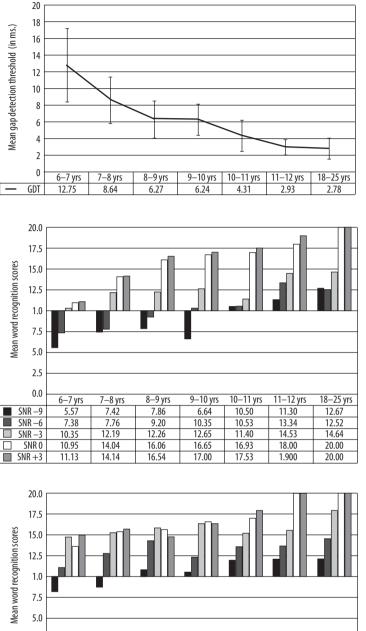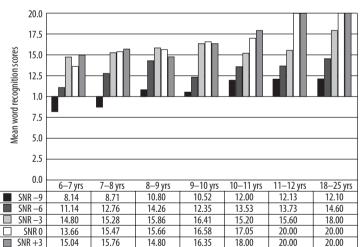
Figure 3. Mean gap detection thresholds (ms, vertical axis) with age (horizontal axis). Smaller values indicate better performance

	6–7 yrs	7–8 yrs	8–9 yrs	9–10 yrs	10–11 yrs	11–12 yrs	18–25 yrs
GDT	12.75	8.64	6.27	6.24	4.31	2.93	2.78

Figure 4. Mean word recognition scores in the presence of continuous noise at various SNR levels (white/black/grey). The inverted bars represent scores below 50% (SNR50)

	6–7 yrs	7–8 yrs	8–9 yrs	9–10 yrs	10–11 yrs	11–12 yrs	18–25 yrs
SNR −9	5.57	7.42	7.86	6.64	10.50	11.30	12.67
SNR −6	7.38	7.76	9.20	10.35	10.53	13.34	12.52
SNR −3	10.35	12.19	12.26	12.65	11.40	14.53	14.64
SNR 0	10.95	14.04	16.06	16.65	16.93	18.00	20.00
SNR +3	11.13	14.14	16.54	17.00	17.53	1.900	20.00

Figure 5. Mean word recognition scores in the presence of interrupted noise at various SNR levels (white/black/grey). The inverted bars represent WRS scores below 50% (SNR50)

	6–7 yrs	7–8 yrs	8–9 yrs	9–10 yrs	10–11 yrs	11–12 yrs	18–25 yrs
SNR −9	8.14	8.71	10.80	10.52	12.00	12.13	12.10
SNR −6	11.14	12.76	14.26	12.35	13.53	13.73	14.60
SNR −3	14.80	15.28	15.86	16.41	15.20	15.60	18.00
SNR 0	13.66	15.47	15.66	16.58	17.05	20.00	20.00
SNR +3	15.04	15.76	14.80	16.35	18.00	20.00	20.00

Word recognition scores in the presence of interrupted noise were better compared to those in the presence of continuous noise. Children as young as 6–7 years performed well in this test and SNR50 was obtained at the −6dB level. Further, comparable adult values in the WRS were observed in the 8–9 year age group where SNR50 was obtained for the −9dB level. These results were far better than those obtained for continuous noise, where this level of performance was reached only after 10–11 years of age.

One-factor analysis of variance (ANOVA) was carried out to estimate the effect of age on word recognition performance. Between-group comparisons at all SNR levels for both continuous and interrupted noise revealed a significant effect of age on word recognition performance (Table 2). Bonferroni post hoc comparison at each SNR level also revealed statistically significant differences between the youngest age group and the adult group. A three-factor mixed ANOVA was conducted to estimate the interaction between word recognition performance and age,

Table 2. Results of one-factor ANOVA estimating the effect of age on word recognition performance

Type of noise	SNR	Sum of squares	Mean square	*F*-value	*p*-value
Continuous	−9 dB	294.371	49.062	11.119	0.001
	−6 dB	223.086	37.181	8.916	0.042
	−3 dB	266.771	44.462	9.932	0.032
	0 dB	321.143	53.524	9.190	0.021
	+3 dB	219.686	36.614	5.728	0.019
Interrupted	−9 dB	385.471	64.245	8.583	0.000
	−6 dB	174.571	29.095	5.984	0.000
	−3 dB	46.443	7.740	2.205	0.046
	0 dB	115.271	19.212	5.237	0.010
	+3 dB	135.500	22.583	6.298	0.007

Table 3. Results of the 3-factor mixed ANOVA showing difference in word recognition scores as a function of age, type of noise, and SNR levels

Source	df	*F*-value	*p*-value
Age	6	27.296	0.000
Type of noise	1	444.038	0.000
SNR	4	23.473	0.026
Age*type of noise	6	2.671	0.016
Age*SNR	24	7.642	0.001
Type of noise*SNR	4	248.530	0.000
Age*type of noise*SNR	24	3.459	0.039

type of noise, and SNR level. As evident from Table 3, the results reveal major effects of age, type of noise, and SNR level on word recognition performance.

Categorical perception of stop consonant

A 13-step voicing continuum was presented and the identification function of the percentage correct response was plotted. The identification function was measured by estimating the percentage correct response of voiced and voiceless sound at each step of the continuum for each participant. This was plotted as an averaged response at each VOT step, and the intersection of the graphs drawn for |ba| and |pa| was considered the categorical boundary. This boundary was the point on the graph where 50% correct responses for either |pa| or |ba| were obtained.

Average identification graphs for each age group are shown in Figure 6. As evident from the figure, all participants consistently responded to |ba| for the first step along the voicing continuum, with a VOT cutoff value of 0 ms, after which the responses gradually shifted to |pa| for the last step at a VOT value of 60 ms. The shift in perception occurred differently for different age groups.

The shift in the categorical boundary was gauged by obtaining the categorical boundary for each participant. The categorical boundary was determined by measuring the

VOT value that results in either a |pa| or |ba| response 50% of the time by using a logistic regression model and a linear or nonlinear interpolation function. An individual participant's responses for all five trials were averaged across each step along the continuum and the response line was fitted with either a linear or nonlinear regression depending on the distribution of the responses using Prism software (version 5.03; GraphPad Software Inc.). Once the line was fitted to the curve, the 50% probability response was interpolated.

The VOT cut off value measured in this way was compared as a function of age using one-way ANOVA. The mean VOT cut off scores at the categorical boundary for 6–7 year old children were shifted more towards the perception of voiced sound in comparison to adults (Figure 7). As evident from Figure 7, there was a gradual change in the mean VOT cut off scores at the categorical boundary, with the boundary shifting more towards the perception of voiceless sound as age increased. The ANOVA results also revealed a major effect of group on the absolute VOT cut off scores at the categorical boundary. Bonferroni post hoc comparison between groups indicated a significant difference between the 6–7 year group and the adult group. However, no such difference was observed between children 9–10 years old and adults. This result indicates that the mean VOT cut off scores at the categorical boundary reach adult-like values by 9–10 years of age. The results
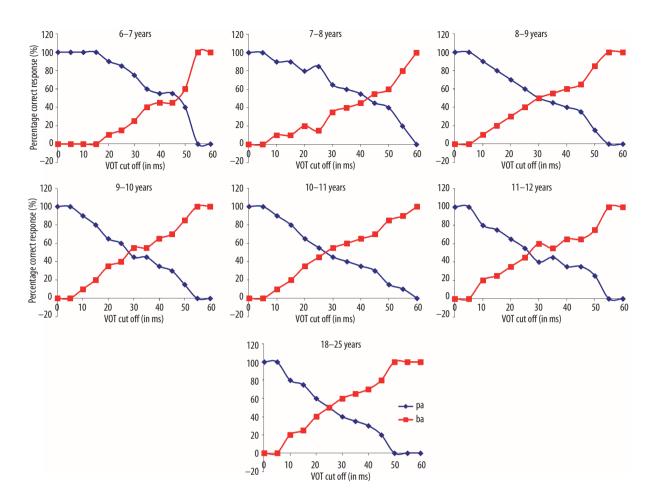
Figure 6. Mean percentage correct responses to |ba| (blue) and |pa| (red) as a function of voice onset time (VOT), producing sets of identification functions. Each panel is for a different age group (marked at top)
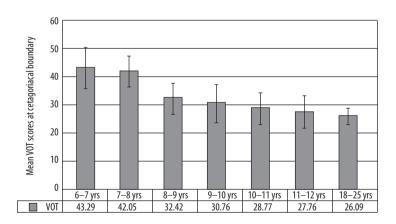

Figure 7. Mean VOT scores at the categorical boundary for different age groups. Error bars represent standard deviations

	6–7 yrs	7–8 yrs	8–9 yrs	9–10 yrs	10–11 yrs	11–12 yrs	18–25 yrs
VOT	43.29	42.05	32.42	30.76	28.77	27.76	26.09

suggest that although categorical perception develops before 6 years of age, the boundary precision continues to develop until 9–10 years and beyond.

Time compressed speech test

The results of the word recognition tests were comparable with those of TCST, in which children of age 6–7 years were able to identify sentences correctly more than 50% of the time for a compression of up to 60%. By 7–8 years of age, 50% correct responses were obtained for compression of up to 70%. Beyond 8 years, no improvement in scores was observed. At 80% time compression, none of the participants were able to identify 50% of the sentences correctly.

Discussion

The present study has demonstrated that temporal processing abilities are immature in children and gradually develop with age. Although the development of various temporal resolution skills begins early, maturation continues even into the second decade of life. This maturation is task-dependent, such that some tasks require less processing skill and mature quickly, whereas other tasks requiring complex processing continue to develop with age. Thus, the results of the present study support the hypothesis that a systematic development in temporal resolution abilities should be observable in children.

This trend in maturation is well supported by previous researchers. Davis and McCroskey [16] found that gap detection abilities reach adult values by 8 years of age, a finding similar to that of Wightman et al. [28] as well as that of the present study. However, some inconsistencies in the literature prompted the present study. Researchers such as Shinn et al. [29] found no significant difference in the gap in noise scores between 7 year and 18 year old children. They found that the gap in noise performance reached adult values by 7 years of age, and suggesting it was a viable tool for evaluating temporal resolution abilities in both children and adults. However, this discrepancy may have been arisen due to the nature of the task: the gap in noise test is less complex and requires limited processing ability in comparison to gap detection skills. Whatever the reason, a discrepancy remains in the literature regarding the age at which gap detection abilities develop in children. This variability becomes more evident with the findings of Irwin et al. [17], who estimate that gap detection ability develops until 11 years of age, after which it reaches adult values. This wide discrepancy in the literature may be attributed to differences in stimuli and participants. Nevertheless, when the results of the majority of studies (including the present study) are considered, it may fairly confidently be stated that gap detection ability reaches adult values by 7–8 years of age.

Similar and even wider inconsistencies in estimating the maturation age of modulation perception have also been reported in the literature. Hall and Grose [30] studied the developmental trajectory of TMTF and found it to reach adult-like values by 4–5 years of age, in contrast to other findings which report that children as old as 8–11 years had significantly higher modulation thresholds at 4 Hz in comparison to somewhat older children and adults [31]. This finding is in agreement with the present study where it was found that modulation abilities reach adult values at 10–11 years of age. Even Hall and Grose, who stated that an adult-like configuration is reached by 4–5 years of age, found that young children were less sensitive than adults in modulation detection. Thus, considering the observation of the present study, together with the findings of Peter et al. [31], it may be more correct to consider 10–11 years as the maturation age of temporal modulation detection. In any case, modulation rate as a factor should not be ignored.

The above results show that, in comparison with gap detection, modulation detection ability matures later. Since both measures of temporal resolution ability involve non-speech

stimuli, a discrepancy indicates there must be a different neural mechanism involved in processing the two types of temporal stimuli. Sakai et al. [32] found that the processing of temporal modulation transfer functions is related to the degree of stimulus locking at the level of the primary auditory cortex. Similarly, the gap in noise test is also sensitive to cortical and brainstem lesions [33], indicating that its neural correlate must reside somewhere in the auditory cortex or higher brainstem. The possible difference may be at the level of neurons, since the TMTF is more sensitive to second-order neural inhibition in the primary auditory cortex [32]; however, for gap detection, no such explanation has been offered in the literature. Nonetheless, involvement of the brainstem along with the auditory cortex in the processing of temporal gaps indicates that the activation is probably more peripheral than in the deep structures of the auditory cortex. In the present study, therefore, the findings may be attributed to the differential maturation rate of the two types of temporal stimuli. Since myelination of the peripheral auditory cortex occurs faster than the deep auditory structures, the inference is that the GDT should mature earlier than the TMTF [34].

This finding may also explain the variation in maturation age with different modulation rates. According to Eggermont [35], the primary auditory cortex is responsible for processing stimuli with high modulation rates, whereas the secondary auditory cortex is responsible for processing auditory stimuli with lower modulation rates. Vaughan and Kurtzberg [36] noted that the primary regions of auditory cortex complete myelination earlier than the secondary auditory cortex; hence, stimuli with higher modulation rates should show an adult-like configuration earlier than stimuli with lower modulation rates. These findings suggest a link between cortical myelination and maturation of temporal resolution ability, although a causal relationship requires objective functional imaging techniques.

The detection of temporal change was surprisingly immature even by 12 years of age. This result may be discussed in terms of the stimuli. Temporal change detection requires identification of a gap in a sequence of five tones, where the gap duration between tone 3–4 or 4–5 was varied adaptively. In the 3AFC procedure, one of the alternatives had a gap whereas the remaining two alternatives did not have any gap. Thus, the complexity of the task was twofold: one had to identify the alternative with a gap, whereas the gap location was also adaptively moving from one tone interval to other. Thus, more attention and higher order processing skills are required to identify the gap appropriately. This may be the probable reason of poor performance of children in this test in comparison to adults. However, no prior research study is available to confirm the results; the finding is novel to the present study.

Irrespective of the findings of the TCD test, the other two findings indicate that temporal resolution abilities develop by around 10–11 years of age, at least for non-speech stimuli. The findings accomplished the aim of the present study: to show a developmental pattern in temporal resolution abilities (at least for non-speech stimuli); here a developmental trend was found which had a variable trajectory depending on the stimuli. These findings correlate with the results of the temporal resolution tests,

which use speech stimuli. A positive correlation was observed between TMTF and word recognition scores, where in both the tests adult values were reached by around 10–11 years of age.

The finding here for the word recognition test has been well discussed. A number of previous investigators have employed speech perception in competing noise to assess temporal resolution abilities [9,37–41]. In one such extensive study by Stuart [6], word recognition scores in children were assessed in the presence of continuous noise. The results of his study revealed significant difference in the mean WRS scores at 0 dB SNR between 10–11 year old children and adults. However, in the present study, at 0 dB SNR, children as young as 6–7 years also performed well with WRS scores above 50%. This difference may be attributed to differences in methodology and in test stimuli.

The better responses for interrupted noise in comparison with continuous noise may be due to the ability of subjects to determine spectral cues between the noise segments [42]. Earlier researchers have also shown poorer performance in WRS in the presence of interrupted noise for children in comparison to adults. Stuart [6] reported that although performance with interrupted noise was better than that with continuous noise, adult values were reached only after 10–11 years of age. Again, the maturation age for the participants in Stuart's study was higher in comparison to that of the participants in the present study, and may be attributed to methodological and stimuli-related variation. The results of the word recognition scores obtained in the present study correlated well with those of the gap detection scores: in both tests, early maturation was observed.

As with the performance in other tests, the findings of categorical perception ability showed a developmental trend. Since both categorical perception and gap detection share the same underlying mechanism [7], it may be justified to include this particular test in a test battery for assessing development related changes in temporal processing abilities. It has already been suggested that infants as young as 2 months display a pattern of discrimination consistent with categorical perception [43]. In another experiment using a high amplitude sucking paradigm for infants 1–4 months old, it was found that infants were able to demonstrate the difference between voiced and voiceless sound when the stimuli was changed in terms of VOT. However, in none of these studies did the researchers report that the categorical boundary (in terms of VOT) for infants was the same as that for adults. In the present study also, it was found that the children as young as 6 years of age were able to identify stimuli categorically, but the categorical boundary for young children was more variable than in older children and adults. This finding is supported by Medina and Serniclaes [44] who found late development of categorical perception of speech sounds in children aged 8–11 years. The researchers also found a steeper slope at the categorical boundary for adults in comparison to children, a finding consistent with the present study.

Hazen and Barrett [45] also reported that the identification of voicing contrast in children was poorer than in adults. They found that perceptual function was significantly

different in 6 year old children compared to 12 year olds. However, according to them, children never achieve adult-like proficiency until 12.6 years of age, contradicting the findings of the present study in which adult-like scores were achieved at 9–10 years of age. An explanation for this discrepancy can be found in the study of Parnell and Amerman [46] who found that although children as young as 11 years are able to identify consonants as well as adult, the children's responses were less consistent and involved more error substitutions. This situation was clearly evident in the present study also, where it was noted that the standard deviation for the VOT scores in the paediatric group was far more than in the adult group (Figure 7). Analysis of data from individual participants revealed that the range of responses was more variable in children than in adults. Higher variability in children may be attributed to immature processing skills, which continue to develop even in the second decade of life.

Finally, the time compressed speech test was administered on children to study the developmental trend. Although the TCST is not a direct measure of temporal resolution, and is more sensitive to auditory closure deficits [8], it does assess the speed of speech perception [9] which may be indirectly related to temporal resolution ability. The findings for the TCST were also found to be similar to those of the other temporal resolution tests. The results showed a strong developmental trend, where the youngest children showed the ability to identify speech with 60% time compression and the values reached adult-like values by 8 years of age. These numbers are well accepted in the research literature. Beasley et al. [47] were the first to try and develop normative data for time compressed speech in children, and found that for PBK-50 words as stimulus, children at 4 years of age achieved less than 50% correct responses at 60% compression, a figure which gradually improved with age. The present study provides information up to a compression level of 80%. Increased compression amplifies the sensitivity of the stimuli to assess temporal resolution.

In summary, we have observed that temporal resolution abilities start developing in the early stages of life, even before 6 years of age, and the development process continues even after 12 years of age. However, the developmental trajectory is different for different stimuli, including both speech and non-speech stimuli. Hence, it is necessary to establish age-specific as well as stimuli-specific norms, a process achieved in the present study.

Conclusions

The present study has focused on the developmental pattern of temporal resolution abilities in children aged 6–12 years, seeking to find a maturational age, i.e. the age at which temporal resolution skill reaches an adult-like value. The results show that temporal resolution abilities generally reach an adult state by 10–11 years of age. However, different tests produce different ages of maturation. These findings may be useful when assessing children's speech perception, and the various maturational ages should be kept in mind before any firm diagnosis is made. The results may also be helpful in planning therapy for children in whom auditory processing abilities appear compromised.

Acknowledgement

Conflict of interest

The authors declare no conflict of interest.

References:

1. Eggermont JJ. Between sound and perception: reviewing the search for a neural code. Hear Res, 2001; 157(1–2): 1–42.
2. Verhey JL. Temporal resolution and temporal integration; in Plack CJ (ed): The Oxford Handbook of Auditory Science: Hearing. Oxford University Press, Oxford, New York, 2011; 3: 105–22.
3. Erber NP. Speech perception by profoundly hearing impaired children. J Speech Hear Disord, 1979; 44(3): 275–9.
4. Rosen S, Walliker J, Brimacombe JA, Edgerton BJ. Prosodic and segmental aspects of speech perception with the House/3M single channel implant. J Speech Hear Res, 1989; 32(1): 93–111.
5. Schneider BA, Pichora-Fuller MK. Age-related changes in temporal processing: implications for speech perception. Semin Hear, 2001; 22(3): 227–40.
6. Stuart A. Development of auditory temporal resolution in school-age children revealed by word recognition in continuous and interrupted noise. Ear Hear, 2005; 26: 78–88.
7. Elangovan S, Stuart A. Natural boundaries in gap detection are related to categorical perception of stop consonants. Ear Hear, 2008; 29(5): 761–74.
8. Keith RW. Standardization of time compressed speech test. J Edu Audiol, 2002; 10: 15–20.
9. Jafari Z, Omidvar S, Jafarloo S. Effect of aging on speed and temporal resolution of speech stimuli in older adults. Med J Islam Repub Iran, 2013; 27(4): 195–203.
10. Bacon SP, Viemeister NF: Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners. Audiology, 1985; 24(2): 117–34.
11. Fitzgibbons PJ, Wightman FL. Gap detection in normal and hearing-impaired listeners. J Acoust Soc Am, 1982; 72(3): 761–5.
12. Ahmmed A, Parker D, Adams C, Newton V. Auditory temporal resolution in children with specific language impairment. J Med Speech Lang Pathol, 2006; 14(2): 79–96.
13. Norrelgen F, Lacerda F, Forssberg H. Temporal resolution of auditory perception and verbal working memory in 15 children with language impairment. J Learn Disabil, 2002; 35(6): 539–45.
14. Chaubet J, Pereira L, Perez AP. Temporal resolution abilities in students with dyslexia and reading and writing disorders. Int Arch Otorhinolaryngol, 2014; 18(2): 146–9.
15. Maxon AB, Hochberg I. Development of psychoacoustic behavior: sensitivity and discrimination. Ear Hear, 1982; 3: 301–8.
16. Davis S, McCroskey R. Auditory fusion in children. Child Dev, 1980; 51(1): 75–80.
17. Irwin RJ, Ball AK, Kay N, Stillman JA, Rosser J. The Development of auditory temporal acuity in children. Child Dev, 1985; 56(3): 614–20.
18. Buss E, Hall JW-III, Grose JH, Dev MB. Development of adult-like performance in backward, simultaneous and forward masking. J Speech Lang Hear Res, 1999; 42(4), 844–9.
19. Grose JH, Hall JW. Co-modulation masking release: is co-modulation sufficient? J Acoust Soc Am, 1993; 93(5): 2896–902.
20. Elliott LL. Discrimination and response bias for CV syllables differing in voice onset time among children and adults. J Acoust Soc Am, 1986; 80(4): 1250–5.
21. Johnson CE. Children's phoneme identification in reverberation and noise. J Speech Lang Hear Res, 2000; 43(1): 144–57.
22. American National Standards Institute. Methods for manual pure tone audiometry (ANSI S3.21-2004). New York, 2009.
23. Grassi M, Soranzo A. MLP: A MATLAB toolbox for rapid and reliable auditory threshold estimation. Behav Res Methods, 2009; 41(1): 20–28.
24. Boersma P: Praat, a system of doing phonetics by computer. Glot Int, 2001; 5(9/10): 341–5.
25. Loizou P. COLEA: A Matlab software tool for speech analysis, 2009. Retrieved August 18, 2014, from *http://www.utdallas.edu/˜loizou/speech/colea.htm.*
26. Audacity Development Team: Audacity. A free digital audio editor, 1.3.12-beta (Unicode) 2011; Retrieved August 18, 2014, from *http://audicity.sourceforge.net.*
27. Raffin MJ, Thornton AR. Confidence levels for differences between speech-discrimination scores: a research note. J Speech Hear Res, 1980; 23(1): 5–18.
28. Wightman F, Allen P, Dolan T, Kistler D, Jamieson D. Temporal resolution in children. Child Dev, 1989; 60: 611–24.
29. Shinn JB, Chermak GD, Musiek FE. GIN (Gap-In-Noise) performance in pediatric population. J Am Acad Audiol, 2009; 20(4): 229–38.
30. Hall JW, Grose JH. Development of temporal resolution in children as measured by the temporal modulation transfer function. J Acoust Soc Am, 1994; 96(1): 150–54.
31. Peter V, Wong K, Narne VK, Sharma M, Purdy MC, McMahon C. Assessing spectral and temporal processing in children and adults using temporal modulation transfer function (TMTF), Iterated Ripple Noise (IRN) perception, and spectral ripple discrimination (SRD). J Am Acad Audiol, 2014; 25(2): 210–18.
32. Sakai M, Chimoto S, Qin L, Safo Y. Neural mechanism of interstimulus interval-dependent response in the primary auditory cortex of awake cats. BMC Neurosci, 2009; 10(10): 1–21.
33. Musiek FE, Shinn JB, Jirsa R, Bamiou DE, Baran JA, Zaidan E. GIN (Gaps-In-Noise) test performance in subjects with confirmed central auditory nervous system involvement. Ear Hear, 2005; 26(6): 608–18.
34. Moore JK. Maturation of human auditory cortex: implications for speech perception. Ann Otol Rhinol Laryngol Suppl, 2002; 189: 7–10.
35. Eggermont JJ. Representation of spectral and temporal sound features in three cortical fields of the cat. Similarities outweigh differences. J Neurophysiol, 1998; 80(5): 2743–64.
36. Vaughan HG, Kurtzberg D. Electrophysiologic indices of human brain maturation and cognitive development; in Gunnar MR, Nelson CA (eds.); Minnesota Symposia on Child Psychology. Erlbaum, Hillsdale, New Jersey, 1992; 1–36.

37. Suter A. Speech recognition in noise by individuals with mild hearing impairment. J Acoust Soc Am, 1985; 78: 887–900.

38. Bronkhorst AW, Plomp R. A clinical test for the assessment of binaural speech perception in noise. Audiology, 1990; 29(5): 275–85.

39. Gordon-Salant S, Fitzgibbons P. Age effects on discrimination of timing in auditory sequence. J Acoust Soc Am, 2004; 116: 1126–34.

40. Larsby B, Hallgren M, Lyxell B, Arlinger S. Cognitive performance and perceived effort in speech processing tasks: effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. Int J Audiol, 2005; 44(3): 131–43.

41. Narne VK. Temporal processing and speech perception in noise by listeners with auditory neuropathy. PlosOne, 2013; 8(2): e55995.

42. Stuart A, Phillips DP. Word recognition in continuous and interrupted broadband noise by young normal-hearing, older normal-hearing, and presbyacusic listeners. Ear Hear, 1996; 17(6): 478–89.

43. Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. Speech perception in infants. Science, 1971; 171(3968): 303–6.

44. Madina V, Serniclaes W. Late development of categorical perception of speech sounds in pre-adolescent children. ZAS Papers in Linguistics, 2005; 42: 13–31.

45. Hazan V, Barrett S. The development of phonemic categorization in children aged 6–12. J Phonetics, 2000; 28: 377–96.

46. Parnell MM, Amerman JD. Maturational influence on the perception of coarticulatory effects. J Speech Hear Res, 1978; 21: 682–701.

47. Beasley DS, Maki JE, Orchik DJ. Children's perception of time-compressed speech on two measures of speech discrimination. J Speech Hear Disord, 1976; 41(2): 216–25.