ISSN: 2083-389X,
eISSN: 2084-3127

J Hear Sci, 2021; 11(1): 31–39
DOI: 10.17430/JHS.2021.11.1.3

AUDITORY STREAM SEGREGATION WITH SINUSOIDALLY AMPLITUDE MODULATED TONAL STIMULI IN INDIVIDUALS WITH SENSORINEURAL HEARING LOSS

Contributions:
A Study design/planning
B Data collection/entry
C Data analysis/statistics
D Data interpretation
E Preparation of manuscript
F Literature analysis/search
G Funds collection

Jawahar Antony P.[A-F], Animesh Barman[A,D-E]

Department of Audiology, All India Institute of speech and Hearing, India

Corresponding author: Jawahar Antony P; Department of Audiology, All India Institute of Speech and Hearing, Manasagangothri, 570006, Mysore, India; email: jawaharantony@gmail.com; Phone: +91 9845615830

Abstract

Aim: To study auditory stream perception in sensorineural hearing loss using sinusoidally amplitude modulated signals (SAM).

Material and methods: There were 30 participants with normal hearing and 30 participants with SNHL who participated in the study. Two experiments were conducted. In Experiment I, an AB sequence of SAM stimuli was presented having a standard or reference modulation frequency in the A stimuli and a comparison or target modulation frequency in the B stimuli. In Experiment II, only a B stimuli sequence was presented. Both a low carrier frequency of 1 kHz and a high carrier frequency of 4 kHz were used in the experiment. A lower standard modulation frequency of 16 Hz and a higher standard modulation frequency of 256 Hz were considered. The comparison modulation frequencies ranged from 1 octave above to 4 octaves above the standard modulation frequencies. The objective listening task was to pick the irregularity in the rhythmic sequence when different levels of time delays were introduced, and this delay was used as a measure of stream perception.

Results: There was a significant difference between the normal hearing group and the SNHL group in their ability to detect irregularities when higher standard modulation frequencies were used, irrespective of the carrier frequencies. The SNHL group identified the irregularities better than the normal hearing group, indicating they had poorer stream perception.

Conclusions: Poorer stream perception with sinusoidal amplitude modulated tonal stimuli in the SNHL group can be attributed to poorer frequency resolution in the SNHL group.

Key words: sensorineural hearing loss • auditory stream segregation • sinusoidal amplitude modulation

SŁUCHOWA SEGREGACJA STRUMIENIA Z UŻYCIEM BODŹCÓW TONALNYCH O AMPLITUDZIE MODULOWANEJ SINUSOIDALNIE U OSÓB Z NIEDOSŁUCHEM ODBIORCZYM

Streszczenie

Cel: Analiza słuchowej percepcji strumienia w niedosłuchu odbiorczym z zastosowaniem sygnałów o amplitudzie modulowanej sinusoidalnie (SAM).

Materiał i metody: W badaniu wzięło udział 30 uczestników z normalnym słuchem i 30 z niedosłuchem odbiorczym. Przeprowadzono dwa eksperymenty. W eksperymencie I uczestnikom prezentowano sekwencję bodźców SAM o konstrukcji AB, gdzie modulacja częstotliwości bodźców A była standardowa lub referencyjna, a bodźców B porównawcza lub docelowa. W eksperymencie II prezentowana była tylko sekwencja bodźców B. W eksperymentach wykorzystano zarówno niską częstotliwość nośną 1 kHz jak i wysoką 4 kHz. Uwzględniono niższą standardową częstotliwość modulującą 16 Hz i wyższą standardową częstotliwość modulującą 256 Hz. Porównawcze częstotliwości modulujące były w przedziale od 1 do 4 oktaw powyżej standardowej częstotliwości modulującej. Zadanie słuchowe polegało na wykryciu nieregularności w sekwencji rytmicznej przy wprowadzanych różnych okresach opóźnienia, które to opóźnienie było miarą percepcji strumienia.

Wyniki: Zaobserwowano istotną różnicę pomiędzy grupą osób ze słuchem normalnym a grupą z niedosłuchem odbiorczym pod względem ich zdolności wykrywania nieregularności, gdy stosowana była wyższa standardowa częstotliwość modulacji, niezależnie od częstotliwości nośnych. Osoby z grupy z niedosłuchem odbiorczym wykrywały nieregularności lepiej, niż z grupy ze słuchem normalnym, co wskazuje na ich gorszą percepcję strumienia.

Wnioski: Gorszą percepcję strumienia w grupie osób z niedosłuchem odbiorczym przy zastosowaniu bodźców tonalnych o amplitudzie modulowanej sinusoidalnie można przypisać gorszej rozdzielczości częstotliwościowej osób z niedosłuchem odbiorczym.

Słowa kluczowe: niedosłuch odbiorczy • słuchowa segregacja strumienia • amplituda modulowana sinusoidalnie

Introduction

Auditory stream segregation is a process in which complex sounds are separated into individual auditory streams [1]. The auditory streams are formed by either associating or segregating the sounds into a range of possible sound sources [1]. The cues for auditory stream segregation have been widely studied [1–4]. Sequential grouping or sequential segregation happens when sounds are grouped together as similar, or segregated as dissimilar, by comparing the acoustic properties of the preceding sounds [5]. One of the major cues for auditory segregation is the frequency separation between two successive sounds [2–6]. Temporal variations like rate and temporal envelope, or just amplitude modulation between successive sounds, have also been found to be cues for stream segregation in individuals with normal hearing [1,7,8].

There are a few reports in the literature where the phenomenon of auditory stream segregation is seen to be affected in individuals with cochlear hearing loss [5,9]. Poor auditory stream segregation has been noticed in patients with cochlear hearing loss even when there is a large frequency separation between the two successive tones [5,9]. This is attributed to poor frequency selectivity in these individuals. However, the literature also suggests that whenever there is a change in the temporal characteristics of successive sounds, there is no significant difference in stream percepts between individuals with normal hearing and those with cochlear hearing loss [10]. Further, many studies on stream perception in individuals with sensorineural hearing loss are inconclusive, based on either spectral or temporal cues [8,12–14].

Environmental sounds are a mixture of both spectral and temporal variations. Sinusoidal amplitude modulation (SAM) signals are sounds in which changes in parameters can provide either temporal or spectral cues or both [5]. A low carrier frequency with higher modulation frequency provides spectral cues, whereas a high carrier frequency with lower modulation frequency provides temporal cues. Perceptual streams are observed when SAM stimuli are used to study stream segregation in individuals with normal hearing. Based on these results, it has been concluded that variations in the carrier frequency, modulation rate, and modulation depth result in stream segregation [5].

To the best of our knowledge there are no reports available on stream segregation using SAM stimuli in individuals with sensorineural hearing loss (SNHL). However, studies of perception of SAM in individuals with cochlear hearing loss show effects of SNHL [14]. Individuals with cochlear hearing loss fail to judge the modulation depth of SAM stimuli as effectively as normal hearing subjects do. Koopman [14] speculated that the perception of SAM in hearing loss may depend on the formation of stream segregation.

The aim of this study was to see whether the phenomenon of stream segregation was different in individuals with SNHL. The study compared stream segregation between individuals with normal hearing and those with SNHL. Because stream segregation depends on both spectral and temporal cues, the current study attempted to examine the utility of both spectral information and temporal information in SAM stimuli to provide stream segregation in individuals with SNHL. Understanding the cues utilized for stream segregation by individuals with hearing impairment may help in designing better noise reduction algorithms for hearing aids to give improved perception of speech in noise.

Material and methods

Subjects

There were 30 normal hearing adults (mean age 27.4 years) and 30 subjects with sensorineural hearing loss (mean age 34.6 years) who participated in the study. The subjects with normal hearing had no otological complaints, and their pure tone thresholds were within normal limits at clinical audiometric test frequencies. The subjects with SNHL had bilateral mild or moderate hearing loss with a flat audiogram configuration. Their speech identification scores were proportional to the degree of hearing loss. All participants had either 'A' or 'As' type tympanograms; the normal hearing group had normal acoustic reflex thresholds and the SNHL group had elevated or absent acoustic reflexes depending on the degree of hearing loss. TEOAEs were present in all the participants in the normal hearing group and absent in the SNHL group. Informed consent was obtained from all participants. Prior to the study, approval was obtained from the bio-behavioural research committee of the institution (WOF-0348/2014-15). The committee verified that the study procedures followed ethical guidelines and checked that the procedures did not present any hazard to the participants.

Generation of stimuli

The SAM signal was generated based on earlier studies [9,15] using AUX Viewer version 1.0 software. The SAM signal had a carrier frequency (f_c) of 1 kHz and a modulation frequency (f_{mod}) of either 16 Hz or 256 Hz. In addition, a SAM signal with an f_c of 4 kHz was generated. Both these stimuli were considered as standard or reference stimuli. For target or comparison stimuli, SAM was generated with f_{mod} values 1, 2, 3, and 4 octaves higher than the f_{mod} of the standard stimuli for the two f_c frequencies. The sampling frequency for generating the stimuli was 44.1 kHz and a 10 ms cosine ramp was used. The modulation depth of the SAM was kept at 100%. The duration of each SAM stimulus was 60 ms. Adobe Audition software (version 3.0) was used to align the SAM stimuli in a sequence as shown in Figure 1.

Procedure

A listening task was used to measure stream segregation. The procedure was similar to the objective listening experiment for stream segregation proposed by Roberts [15]. Two experiments were conducted on both groups to measure stream perception, and in each a standard sequence and a target sequence were used to measure stream segregation.

Experiment I

In the first experiment, two AB sequences, a standard and a target sequence of SAM stimuli were presented (Figure 1a,b). There were 12 pairs of AB stimuli in each sequence. In the *standard* AB sequence, the A stimuli had a constant f_c and f_{mod}. The B stimuli in the AB sequence also had the same f_c as that of the A stimuli but the f_{mod} was in a higher octave (1, 2, 3, or 4 octaves) to the f_{mod} of the A stimuli. Two f_c and two f_{mod} combinations were used as A stimuli in the AB sequence. Similarly, for the B stimuli, two f_c values and four f_{mod} combinations were used. The standard sequence had a constant gap of 40 ms between the two AB stimuli in the AB sequence (Figure 1a). But for the target sequence, although the delay between the first six pairs was constant at 40 ms, from the seventh pair, the delay from B to the next A was increased by 8 ms (ΔT). Then the delay increased in the order 2 ΔT, 3 ΔT, and 4 ΔT for the next four pairs, after which the delay stayed at 4 ΔT for the last two pairs (Figure 1b). When the delay was increased within an AB pair, the silent period between the next AB pair was adjusted to keep the overall duration of the sequence constant. The cumulative delay (4 ΔT) for the initial target sequence was kept at 32 ms. The value of 32 ms was steadily reduced in each run by a factor of 1.189 until the subject could not identify the irregularity among the given two sequences. Then the step-size was increased (called a reversal) until the subject could again identify the irregularity in one of the sequences. This process was repeated until a psychometric curve could be produced from which a threshold cumulative delay could be obtained.

This progressive increase in delay in the target sequence caused the subject to hear an irregular or arrhythmic sequence [15]. The experiment was conducted through Matlab (version R2014a) via a personal computer (Sony Vaio model SVE14125). The output was routed through a calibrated audiometer (Inventis Piano, Italy) to HDA 200 headphones. The standard and target sequence were presented diotically to the subjects at their most comfortable levels. A two alternative forced choice method (2 AFC) was used. The subjects were instructed to find the arrhythmic sequence out of the two given sequences. The minimum cumulative delay (*d1*) was calculated as the 70.7% point in the psychometric function, as given by Levitt [16]. The results of the experiment determine the level at which the subject can detect irregularity in the target sequence. The formation of separate A and B streams from the AB sequence makes it difficult to detect the irregularity, resulting in a larger *d1* [15].

Experiment II

Experiment II also had a standard sequence and a target sequence. In this experiment, only the SAM of the B stimulus in the AB sequence used in Experiment 1 was employed. Hence, instead of 12 pairs, 24 stimuli arranged in sequence were used. The standard sequence had 24 stimuli with equal delay between each stimulus (Figure 1c). In the target sequence, the first 13 stimuli had equal intervals, but from the 14th stimulus on, a delay of ΔT was introduced (similarly to Experiment I). Likewise, as in that experiment, a progressive delay was introduced in 16th, 18th, and 20th stimulus by 2 ΔT, 3 ΔT, and 4 ΔT respectively. The 22nd and 24th stimulus had a delay of 4 ΔT (Figure 1d). The rest of the procedures were similar to Experiment I, and subjects were instructed to pick the arrhythmic sequence out of the two given sequences. The minimum cumulative delay (*d2*) for Experiment II was obtained as in Experiment I. The cumulative delay (4 ΔT) for the initial target sequence was begun at 32 ms, and the step size of the cumulative delay was reduced by a factor of 1.189 for each reversal as in Experiment I. Since there was only a B sequence, no stream segregation was possible and hence the *d2* was the minimum cumulative delay without any stream segregation. This *d2* served as the reference.

The *d1* obtained for each stimulus from Experiment I were subtracted from the *d2* obtained for corresponding stimuli from Experiment II and denoted as the difference (*D*). The *D* values give the quantum of the stream percept. The *d1*, *d2*, and *D* values were subjected to statistical analysis.

The *d1*, *d2*, and *D* values from the two experiments were analyzed using SPSS (Version 20). A Shapiro–Wilk test of normality showed that the data were not normally distributed ($p < 0.01$). Thus, non-parametric tests were used to compare the *d1*, *d2 and D* values within and across the two groups.

Results

The means and standard deviations (SD) of *d1*, *d2*, and *D* were obtained for both groups as shown in Figures 2, 3, and 4 respectively.

In Figure 2, it can be seen that, for both groups, there are similar *d1* values for low standard f_{mod}, whereas larger differences are seen at higher standard f_{mod}, irrespective of f_c. Subjects with SNHL required a lesser delay in the AB sequence to detect irregularities when a high f_{mod} was used.

Figure 3 shows that in the two groups the *d2* values were similar across all target f_{mod} and f_c frequencies. This indicates that the perception of irregularity in a sequence was similar across the two groups.

A Friedman test was done to see the significance of the difference in *d1* of the normal hearing group for changes in f_c and the standard f_{mod}. A significant main effect was found within the normal hearing group for f_c of 1 kHz with 16 Hz as the standard f_{mod} ($\chi^2(3) = 51.15$, $p < 0.001$, $W = 0.56$); for f_c of 1 kHz with 256 Hz as the standard f_{mod} ($\chi^2(3) = 8.86$, $p = 0.03$, $W = 0.09$); and for f_c of 4 kHz with 256 Hz as the standard f_{mod} ($\chi^2(3) = 37.29$, $p < 0.001$, $W = 0.41$). No main effect was seen for f_c of 4 kHz with 16 Hz as the standard f_{mod} ($p = 0.08$).

Pairwise comparisons were done using a Wilcoxon signed rank test. Table 1 presents the statistical significance between f_{mod} values for the normal hearing group. It can be noted that there is a significant difference in *d1* between the 16 Hz standard f_{mod} and the 256 Hz target f_{mod} for an f_c of 1 kHz in the normal hearing group. It can also be noted that between the 256 Hz standard f_{mod} and most of the target f_{mod} values there were significant differences, regardless of f_c.

Further, Friedman test was done to see the significance of the difference in *d1* values of the SNHL group across carrier frequencies and modulation frequencies. No significant main effect was found within the SNHL group for an f_c of 1 kHz with 16 Hz as the standard f_{mod} ($p = 0.39$); for an f_c of 1 kHz with 256 Hz as the standard f_{mod} ($p = 0.17$); and for an f_c of 4 kHz with 16 Hz as the standard f_{mod} ($p = 0.78$) and 256 Hz as the standard f_{mod} ($p = 0.38$).

A Mann–Whitney *U*-test was done to see the significance of the differences in *d1* values between the normal hearing group and the SNHL group for all f_c and f_{mod}. It was found that there was a significant difference between the

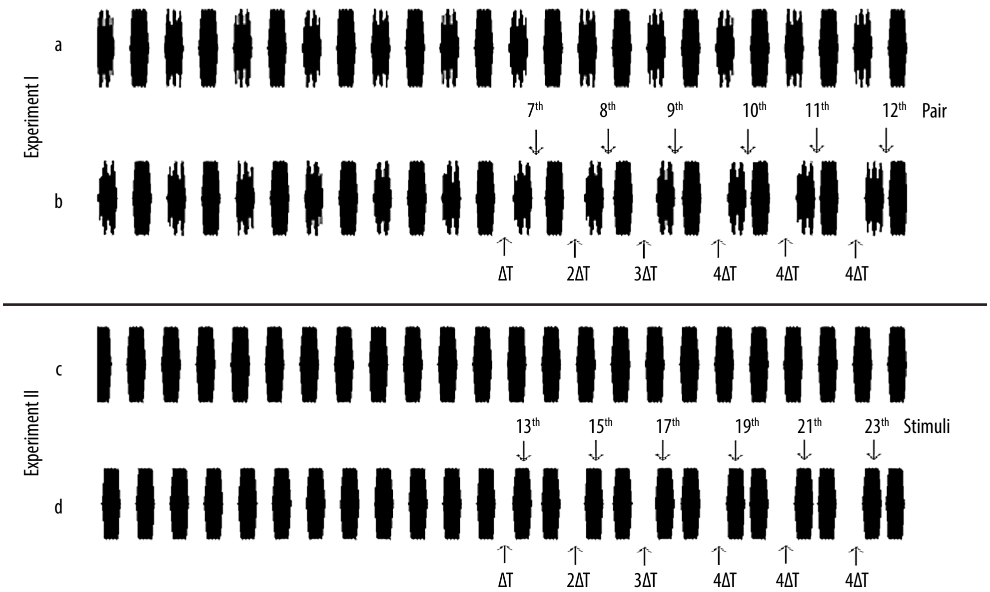

Figure 1. Part **a** represents the *standard* AB sequence and part **b** represents the *target* AB sequence presented in Experiment I. The standard stimuli comprises A of 4000 Hz modulated by 256 Hz and B of 4000 Hz modulated by 512 Hz. The AB sequence of SAM stimuli has an equal interval between the two. However, in the *target* sequence of the AB cycle, the silence between the A and B stimuli is delayed by 8 ms at the 7th cycle (ΔT), 16 ms at the 8th cycle (2 ΔT), 24 ms at the 9th cycle (3 ΔT), and by 32 ms at the 10th, 11th, and 12th cycles (4 ΔT). Part **c** represents the *standard* sequence and part **d** the *target* B sequence of experiment II. Here there are only B stimuli (4000 Hz modulated by 512 Hz) with equal intervals between them in the *standard* stimuli. In the *target* sequence, however, the silent interval between the B stimuli is delayed by 8 ms at the 13th stimuli (ΔT), 16 ms at the 15th (2 ΔT), 24 ms at the 17th (3 ΔT), and by 32 ms at the 19th, 21st, and 23rd stimuli (4 ΔT).

Table 1. Significance in Wilcoxon signed rank test for *d1* between f_{mod} values for the normal hearing group

***f*c**	**Standard *f*mod**	**16 Hz**			**Standard *f*mod**	**256 Hz**		
	Target *f*mod	**64 Hz**	**128 Hz**	**256 Hz**	**Target *f*mod**	**1024 Hz**	**2048 Hz**	**4096 Hz**
1 kHz	32 Hz	*	*	*	512 Hz	NS	NS	*
	64 Hz		NS	*	1024 Hz		NS	*
	128 Hz			*	2048 Hz			NS
4 kHz					512 Hz	*	*	*
					1024 Hz		NS	NS
					2048 Hz			NS

* $p < 0.05$; NS – not significant

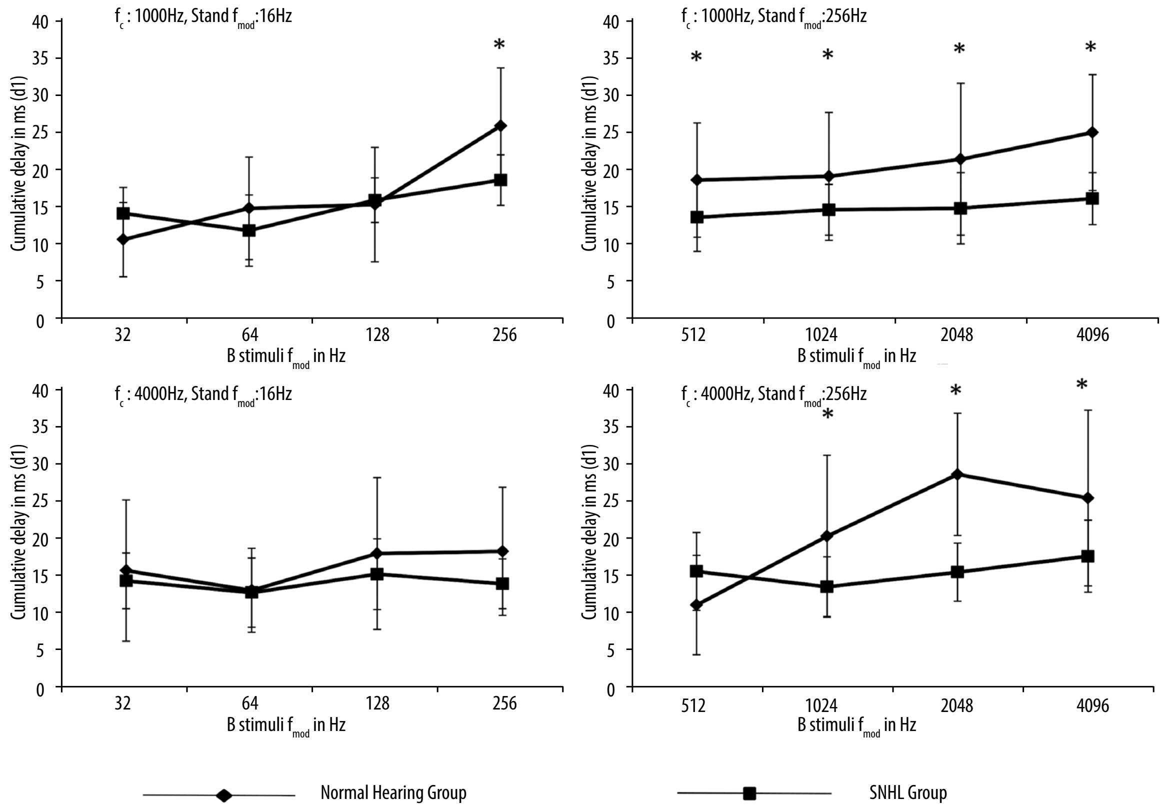

Figure 2. Mean and SD of the *d1* values for the normal hearing and SNHL groups. The asterisks indicate a significant difference ($p < 0.05$) between the normal and SNHL groups

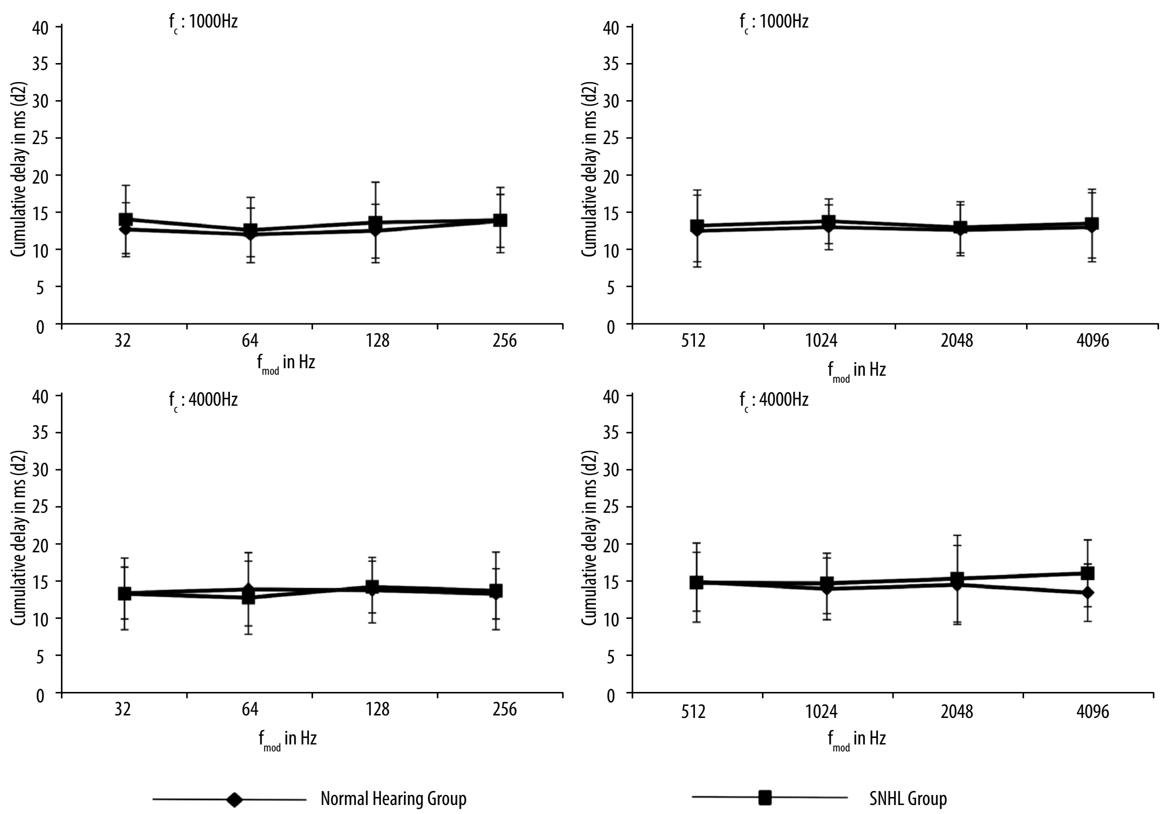

Figure 3. Mean and SD of *d2* values for the normal hearing and SNHL groups

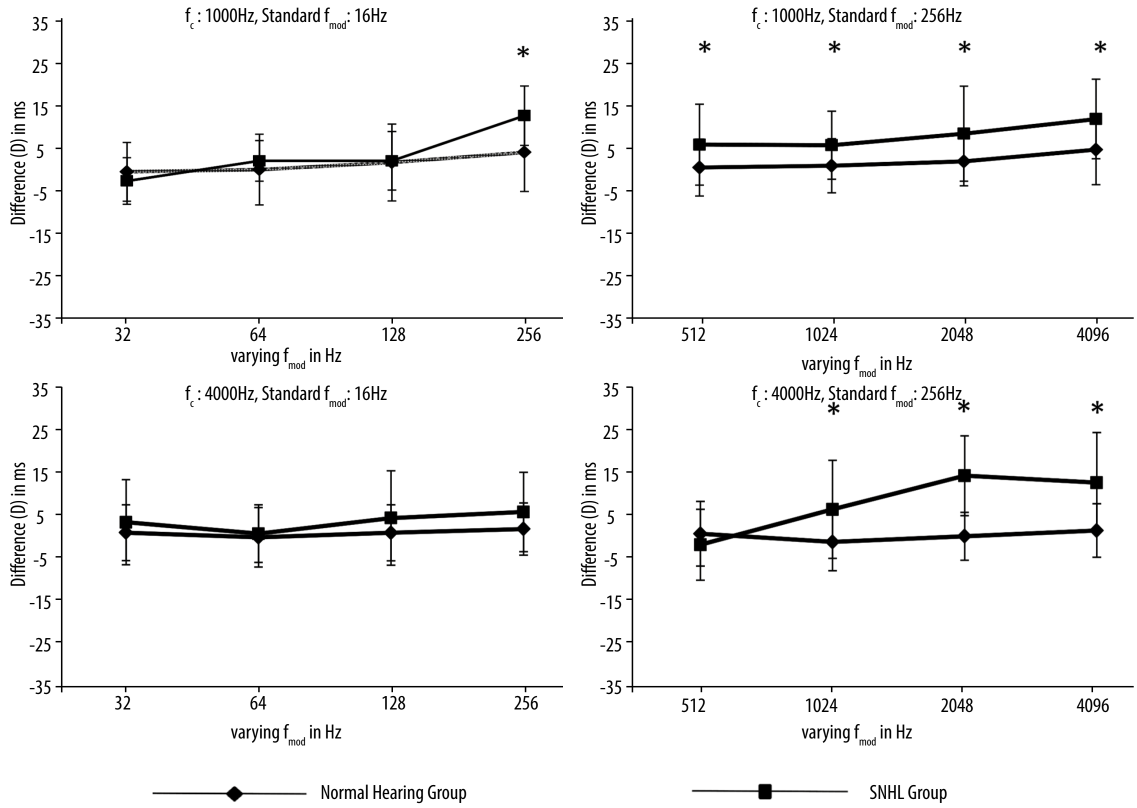

Figure 4. Mean and SD of *D* values for the normal hearing and SNHL groups. The asterisks indicate a significant difference ($p < 0.05$) between the groups

Table 2. Significance in Wilcoxon signed rank test for *D* between the *f*mod values for the normal hearing group

fc	**Standard *f*mod**	**16 Hz**			**Standard *f*mod**	**256 Hz**		
	Target *f*mod	**64 Hz**	**128 Hz**	**256 Hz**	**Target *f*mod**	**1024 Hz**	**2048 Hz**	**4096 Hz**
1 kHz	32 Hz	*	*	*	512 Hz	NS	NS	*
	64 Hz		NS	*	1024 Hz		NS	*
	128 Hz			*	2048 Hz			NS
4 kHz					512 Hz	*	*	*
					1024 Hz		*	NS
					2048 Hz			NS

* $p < 0.05$; NS – not significant

two groups as seen in the Figure 2, where we see a significant difference in *d1* for higher f_{mod} between the normal hearing group and the SNHL group.

A Friedman test was done to see the significance of the difference in *d2* values within each group and across target modulation frequencies and carrier frequencies. No significant main effect was seen for an f_c of 1 kHz with 16 Hz as the standard f_{mod} ($p = 0.36$); for an f_c of 1 kHz with 256 Hz as the standard f_{mod} ($p = 0.17$); for f_c of 4 kHz with 16 Hz as the standard f_{mod}; and for an f_c of 4 kHz with 256 Hz as the standard f_{mod} ($p = 0.35$). Similarly, for the SNHL group, no significant main effect was seen for an f_c of 1 kHz with 16 Hz as the standard f_{mod} ($p = 0.56$); for an f_c of 1 kHz with 256 Hz as the standard f_{mod} ($p = 0.50$); for an f_c of 4 kHz with 16 Hz as the standard f_{mod} ($p = 0.07$); and for an f_c of 4 kHz with 256 Hz as the standard f_{mod} ($p = 0.27$).

The mean and SD of the *D* values obtained by subtracting *d1* and *d2* are shown in the Figure 4. A Friedman test was done to see the significance of the difference in *D* values within each group across the different target f_{mod} and f_c. A significant main effect was found within the normal hearing group for an f_c of 1 kHz with 16 Hz as the standard f_{mod}

($\chi^2(3) = 38.89$, $p < 0.001$, $W = 0.43$); for an f_c of 1 kHz with 256 Hz as the standard f_{mod} ($\chi^2(3) = 8.13$, $p = 0.04$, $W = 0.08$); and for an f_c of 4 kHz with 256 Hz as the standard f_{mod} ($\chi^2(3) = 19.41$, $p < 0.001$, $W = 0.21$). No significant main effect was seen for an f_c of 4 kHz and 16 Hz standard f_{mod} ($p = 0.07$).

Figure 4 shows similar *D* values for the 16 Hz standard f_{mod} for both groups, whereas larger differences are seen with 256 Hz standard f_{mod} irrespective of f_c. This pattern is similar to that observed in the *d1* values.

Pairwise comparisons were done for the normal hearing group using a Wilcoxon signed rank test. Table 2 presents the statistical significance between modulation frequencies for the normal hearing group. It can be noted that few significant differences in *D* occur at the 256 Hz standard f_{mod} in both the f_c.

Within the SNHL group, no significant main effect was found on the *D* values for an f_c of 1 kHz with 16 Hz as the standard f_{mod} ($p = 0.38$); for an f_c of 1 kHz with 256 Hz as the standard f_{mod} ($p = 0.19$); and for an f_c of 4 kHz with 16 Hz as the standard f_{mod} ($p = 0.77$) or for 256 Hz as the standard f_{mod} ($p = 0.39$).

A Mann–Whitney *U*-test was done to see the significance of the differences in *D* values between the normal hearing group and the SNHL group. A significant difference in *D* was noticed between the 16 Hz standard f_{mod} and the 256 Hz target f_{mod} for an f_c of 1 kHz in the between normal hearing group and SNHL group. There were significant differences between the 256 Hz standard f_{mod} and most of the target f_{mod} values, irrespective of f_c.

Discussion

For the normal hearing group, there was a significant difference in *d1* for low modulation frequencies of 1 kHz f_c, indicating that the participants failed to detect the irregular pattern in the AB sequence. This poor irregularity detection could be due to the formation of separate streams for the A and B stimuli [15]. The perception of the A and B sequences as independent makes it difficult for subjects to perceive the irregularity. Thus, the minimum cumulative delay was found to be higher whenever there was formation of two separate streams. The difference was seen even when there was an increase of 1 octave in the B stimuli. These results were similar to those of Dollezal [13] who found that a difference of about 1 to 2 octaves between the A and B tones in an ABA sequence was required to form streams when a low modulation frequency was used at 1 kHz f_c. These results can be attributed to the contribution of both temporal and spectral cues. In the current study, the stream perception with a small difference in f_{mod} of about 1 to 2 octaves between the A and B stimuli at 1 kHz may be due to temporal cues. However, stream perception for larger differences of f_{mod}, of around 3 to 4 octaves between the A and B stimuli at 1 kHz, could be due to spectral cues [15].

There was no significant difference of *d1* at low modulation frequencies with f_c of 4 kHz in the normal hearing group. This is in contradiction to the findings of Dollezal [13], where streams were formed at low f_{mod} for f_c of 4 kHz in normal hearing subjects. This disparity could be due to the methods used to study stream perception. Dollezal [13] used a subjective rating scale where participants were asked to rate whether the given sequence was perceived as a single stream or two streams. In the present study, an objective listening task was used where the participants were asked to identify an irregular rhythmic pattern from which stream perception was calculated. Thus, in this method the participants' ability to identify changes in the gaps between the AB pair is studied. Hence, the results suggest that the cues provided for higher f_c may not be sufficient to produce stream segregation, which could alter the perception of delay in the AB sequence.

In the SNHL group there was no significant difference in *d1* noticed at an f_c of 1 kHz with low f_{mod}. It has been reported that frequency resolution is affected in SNHL, and this would reduce the spectral cues available to perceive sound [17,18]. Spectral cues play an important role in stream segregation in individuals with normal hearing [2,3]. When the difference in frequency of alternating sounds do not overlap in the excitation pattern of cochlea, then there is a higher possibility that individuals with normal hearing will perceive stream segregation [3]. Since frequency resolution is poorer in the SNHL group, the A and B stimuli in the AB sequence may overlap in the excitation pattern of cochlea, resulting in poor stream segregation. This could be the reason that the SNHL group perceive the AB sequence as a single stream and so identify irregularities in the AB sequence better than do individuals with normal hearing sensitivity. In this way, better *d1* values were obtained in the SNHL group. As for temporal resolution, many reports suggest that this is not affected in individuals with SNHL [17]. The findings of our study support the view that spectral cues are more important for stream segregation.

For the normal hearing group there was a significant difference in *d1* for high modulation frequencies at f_c of 1 kHz and 4 kHz. The difference was noticed even when there was a 1 octave difference between the A and B tones, a result in accordance with Dollezal [13] who reported that the predicted spectral difference required to form a stream was higher than the actual results obtained. Hence, Dollezal suggested that both the temporal and spectral cues in a SAM sequence with high f_{mod} contribute to stream perception in normal hearing subjects [13].

For the SNHL group there was no significant difference in *d1* values observed for high modulation frequencies at f_c of 1 kHz and 4 kHz. This indicates that for this group there was no stream formation with higher f_{mod}. This could be because cochlear hearing loss gives rise to poor phase locking ability at high frequencies and poor frequency resolution, which could lead to poor stream segregation [17–20]. Impaired stream segregation at higher f_{mod} due to poor frequency selectivity could be the reason that SNHL produces poorer speech understanding in the presence of noise. Studies have reported that temporal fine structure cues (the high f_{mod}'s in the speech signal) are important for the perception of speech in noise [21]. Hence, these results suggest that there is a need to study the relationship between speech understanding in noise and auditory stream segregation using high f_{mod} in individuals with SNHL.

There was also a significant difference seen in *d1* values at f_{mod} of 256 Hz using f_c of 1 kHz between the normal hearing and SNHL groups. Dollezal [19] suggested that a B stimuli f_{mod} difference of 1–2 octaves with reference to a low standard f_{mod} in the AB sequence may primarily give temporal cues. In comparison, a higher B stimuli f_{mod} difference of 3–4 octaves may provide the spectral cues in a normal hearing group [19]. This could be the reason for the differences seen here between the two groups at this f_{mod}. Since SNHL individuals have poor frequency selectivity, the spectral cues responsible for the higher *d1* values in the normal hearing group could not have contributed to the *d1* in the SNHL group. The absence of stream formation with the AB sequence resulted in better detection of irregularities in the SNHL group [17]. These results again confirm the importance of spectral cues for stream segregation.

A significant difference was also noticed for *d1* for most of the higher f_{mod} at f_c of 1 and 4 kHz between the normal hearing group and the SNHL group. The higher *d1* thresholds in the normal hearing group could have led to these differences. Because the perception of spectral cues and temporal cues of SAM stimuli are intact in the normal hearing group, the two streams of A and B stimuli in the AB sequence could be perceived. This could have led to the normal hearing participants detecting the irregularities poorly. A poor phase-locking ability at high frequencies and poor frequency resolution is reported in SNHL subjects [17–20,22], and so the cues responsible for perceiving two streams in the normal hearing group may not have been available to the SNHL group. This would have led to the formation of a single stream and hence the irregularities in the AB sequence were easily perceived in the SNHL group. In other words, the *d1* values were better in the SNHL group, resulting in a significant difference between the groups. These findings suggest that temporal resolution, as in gap detection, is intact in these individuals. The results of the current study also indicate that temporal cues are less affected in individuals with SNHL. However, spectral cues are vital for stream segregation in individuals with SNHL. Since spectral cues degrade in the presence of noise, this could be the reason for the poor speech perception in noise in these individuals.

There was no significant difference in *d2* between the normal hearing group and the SNHL group for both f_c values across different target f_{mod}. These results indicate that the perception of arhythmicity associated with a change in the delay between the SAM stimuli could be easily identified by the participants. This might be because presentation of the B sequence alone resulted in the formation of a single stream.

There was no significant difference in *d2* between the normal hearing and SNHL groups for all four f_{mod} and for both f_c. This indicates that similar rhythmic perception is seen in both groups whenever the adjacent stimuli are similar. This is because there was no formation of two streams in the B-only sequence in Experiment II. Hence, the irregularities in the sequence were easily identified in both groups. The results also suggest that the SNHL group could detect the irregularities as well as the normal hearing group, since temporal resolution is not typically affected in SNHL individuals [23].

The significant differences in *D* values, seen across the target f_{mod} for all f_c, in the normal hearing group and the SNHL group, was similar to that for *d1* differences. This is probably due to differences seen in *d1* but no differences seen in *d2* for both groups. The larger *D* values represent an increase in the identification of the minimum cumulative delay between the A and B stimuli when an AB sequence is presented. This was observed for both groups. There was a significant difference seen in *D* values for 256 Hz f_{mod} at f_c of 1 kHz between the normal hearing and SNHL groups. This could be due to the poor frequency selectivity in the SNHL group, resulting in better *d1* values and leading to bigger differences in *D* values in the SNHL group [17]. There were also significant differences seen between the normal hearing group and the SNHL group with most of the higher target f_{mod} values at f_c of 1 and 4 kHz. This could be due to poor phase-locking ability at high frequencies and poor frequency resolution in the SNHL group which resulted in better *d1* thresholds, producing differences in *D* values in this group [17–21].

The overall results suggest that spectral cues in SAM stimuli are more effective for stream segregation than temporal cues in individuals with normal hearing. Individuals with SNHL fail to use spectral cues for stream segregation. This can be attributed to the fact that frequency selectivity is affected in these individuals. However, more studies are required to evaluate stream segregation with high frequency spectral cues in individuals with SNHL. These results might help in the development of complex spectral subtraction based noise reduction algorithms for hearing aids.

Conclusion

The results of the current study suggest that the increase in *d1* values in the normal hearing group – indicating the inability to detect irregularities in an AB sequence – are due to the formation of two streams. In normal hearing individuals both temporal and spectral cues play a major role in stream formation. Higher standard f_{mod} values, irrespective of f_c, produced larger stream segregation in the normal hearing group. The formation of two streams was affected in the SNHL group: the *d1* values were better, indicating the formation of only a single stream. The poorer stream perception in the SNHL group is likely due to reduced frequency selectivity and poorer phase locking. We suggest that the higher standard f_{mod} of SAM stimuli could be used to assess the stream perceptual ability of individuals with SNHL.

Acknowledgement

We acknowledge the All India Institute of Speech and Hearing, Mysuru, for providing the infrastructure and resources to conduct the study. We also thank the participants for their co-operation.

References

1. Bregman AS. Auditory Scene Analysis: The perceptual organization of sound. MIT Press; 1994.
2. Bregman AS, Campbell J. Primary auditory stream segregation and perception of order in rapid sequences of tones. J Exp Psychol, 1971; 89(2): 244.
3. van Noorden LP. Minimum differences of level and frequency for perceptual fission of tone sequences ABAB. J Acoust Soc Am, 1977; 61(4): 1041–5.
4. Moore BC, Gockel H. Factors influencing sequential stream segregation. Acta Acustica United with Acustica, 2002; 88(3): 320–33.
5. Rose MM, Moore BC. Perceptual grouping of tone sequences by normally hearing and hearing-impaired listeners. J Acoust Soc Am, 1997; 102(3): 1768–78.
6. Rose MM, Moore BC. Effects of frequency and level on auditory stream segregation. J Acoust Soc Am, 2000; 108(3): 1209–14.
7. Hong RS, Turner CW. Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients. J Acoust Soc Am, 2006; 120(1): 360–74.
8. Grimault N, Bacon SP, Micheyl C. Auditory stream segregation on the basis of amplitude-modulation rate. J Acoust Soc Am, 2002; 111(3): 1340–8.
9. Bayat A, Farhadi M, Pourbakht A, Sadjedi H, Emamdjomeh H, Kamali M, Mirmomeni G. A comparison of auditory perception in hearing-impaired and normal-hearing listeners: an auditory scene analysis study. Iran Red Cres Med J, 2013; 15(11).
10. Mackersie CL, Prida TL, Stiles D. The role of sequential stream segregation and frequency selectivity in the perception of simultaneous sentences by listeners with sensorineural hearing loss. J Speech Language Hear Res, 2001; 44(1): 19–28.
11. Mackersie CL. Talker separation and sequential stream segregation in listeners with hearing loss. J Speech Language Hear Res, 2003; 46(4): 912–8.
12. Stainsby TH, Moore BC, Glasberg BR. Auditory streaming based on temporal structure in hearing-impaired listeners. Hear Res, 2004; 192(1-2): 119–30.
13. Dolležal LV, Beutelmann R, Klump GM. Stream segregation in the perception of sinusoidally amplitude-modulated tones. PLoS One, 2012; 127(9): e43615.
14. Koopman J. On the perception of sinusoidally amplitude modulated signals and its relevance to listening in noise (Doctoral dissertation, Universiteit van Amsterdam), 2004.
15. Roberts B, Glasberg BR, Moore BC. Primitive stream segregation of tone sequences without differences in fundamental frequency or passband. J Acoust Soc Am, 2002; 112(5): 2074–85.
16. Levitt H. Transformed up–down methods in psychoacoustics. J Acoust Soc Am, 1971; 49(2): 467–477.
17. Moore BC. Cochlear hearing loss: physiological, psychological and technical issues. John Wiley & Sons; 2007.
18. Florentine M, Buus S, Scharf B, Zwicker E. Frequency selectivity in normally-hearing and hearing impaired observers. J Speech Hear Res, 1980; 23(3): 646–69.
19. Dubno JR, Dirks DD. Auditory filter characteristics and consonant recognition for hearing-impaired listeners. J Acoust Soc Am, 1989; 85(4): 1666–75.
20. Leek MR, Summers V. Auditory filter shapes of normal-hearing and hearing-impaired listeners in continuous broadband noise. J Acoust Soc Am, 1993; 94(6): 3127–37.
21. Steinmetzger K, Rosen S. The role of periodicity in perceiving speech in quiet and in background noise. J Acoust Soc Am, 2011; 138(6): 3586–99.
22. Dolležal LV, Brechmann A, Klump GM, Deike S. Evaluating auditory stream segregation of SAM tone sequences by subjective and objective psycho-acoustical tasks, and brain activity. Front Neurosci, 2014; 119(8): 1–15.
23. Reed CM, Braida LD, Zurek PM. Review of the literature on temporal resolution in listeners with cochlear hearing impairment: a critical assessment of the role of suprathreshold deficits. Trends Amplif, 2009; 13(1): 4–43.