applied sciences

MDPI

Article

Relationship between Effective Duration and Intelligibility of Japanese Monosyllables in Individuals with Sensorineural Hearing Loss

Sakie Akasaka [1], Tadashi Nishimura [1,*], Ryota Shimokura [2], Tadashi Kitahara [1] and Hiroshi Hosoi [3]

[1] Department of Otolaryngology-Head and Neck Surgery, Nara Medical University, 840 Shijo-Cho, Kashihara 634-8521, Nara, Japan
[2] Graduate School of Engineering Science, Osaka University, 2-1 Yamadaoka, Suita 565-0871, Osaka, Japan
[3] MBT (Medicine-Based Town) Institute, Nara Medical University, 840 Shijo-Cho, Kashihara 634-8521, Nara, Japan
* Correspondence: t-nishim@naramed-u.ac.jp; Tel.: +81-744-22-3051

Abstract: Among the temporal elements in the autocorrelation function, the effective duration (τ_e) is a useful indicator of speech recognition for patients with sensorineural hearing impairment. We assessed the influence of speech recognition performance on the relationship between the percentage of accurately perceived articulation and the median τ_e (τ_e-med) and the relationship between monosyllabic confusion and the τ_e-med. Significant correlations were observed between the articulation percentage and the average τ_e-med in groups with high, middle, and low speech recognition scores (SRSs). Two-factor mixed analysis of variance revealed significant main effects for the condition (presentation/response). There was no significant main effect for group (high-, middle-, or low-SRS) scores and no significant interaction between the groups. The average τ_e-med of the response was significantly longer than that of the presentation in all three groups. Monosyllables with short τ_e-med values tended to be misheard as monosyllables with a long τ_e-med when confusion occurred. The τ_e-med was useful for estimating monosyllables that patients with sensorineural hearing impairment find easy to listen to, independent of speech recognition performance.

Keywords: sensorineural hearing loss; median effective duration (τ_e); speech recognition score

check for updates

Citation: Akasaka, S.; Nishimura, T.; Shimokura, R.; Kitahara, T.; Hosoi, H. Relationship between Effective Duration and Intelligibility of Japanese Monosyllables in Individuals with Sensorineural Hearing Loss. *Appl. Sci.* **2023**, *13*, 8244. https://doi.org/10.3390/app13148244

Academic Editor: Masayuki Takada

Received: 7 June 2023
Revised: 11 July 2023
Accepted: 14 July 2023
Published: 16 July 2023

1. Introduction

In general, aging is associated with clinically moderate sensorineural hearing loss, which can lead to difficulties in understanding speech. Previous studies evaluated the role of temporal, frequency, and pitch information as cues for speech recognition [1,2]. In individuals with sensorineural hearing impairment, frequency resolution and selectivity tend to decrease, contributing to reduced speech recognition [3,4]. At Japanese medical institutions, speech perception in patients with clinically moderate sensorineural hearing loss is evaluated according to the percentage of correctly perceived monosyllables out of a total of 50 Japanese monosyllables, as shown in Table 1. Previous studies investigated speech intelligibility in patients with sensorineural hearing loss and identified the less discernible consonants among the Japanese monosyllables [5,6]. However, to the best of our knowledge, only a few studies explained the differences in discernibility in terms of the physical characteristics of monosyllables. In a previous study, the relationship between speech recognition and the median τ_e (τ_e-med) was analyzed in patients with sensorineural hearing impairment [7]. Shimokura et al. showed that the τ_e-med was strongly correlated with the percentage of correctly perceived articulations of consonants [7].

The autocorrelation function (ACF) is an established indicator for the temporal analysis of auditory nerve processes [8]. Neural processing approximating the ACF occurs in the inferior colliculus, which primarily works on pitch perception [9,10]. Auditory perceptions are strongly related to the timing of nerve firings caused by binaurally detected sounds, and the ACF is modeled on the processors of the auditory nerve [11,12]. Some parameters can be calculated from ACF analyses of monosyllables. These ACF parameters are frequently used for music and acoustics in concert halls [13] and for the evaluation of several types of noise [12–18]. Some studies analyzed the factors of ACF in relation to speech intelligibility in people with normal hearing in different sound-field transmissions [19,20]. Among the temporal elements of ACF, the effective duration (τ_e) reflects the stability of the temporal pattern of sound waves [21]. As mentioned, Shimokura et al. [7] analyzed the relationship between speech recognition and the median τ_e (τ_e-med) in patients with sensorineural hearing impairment and showed that the τ_e-med was strongly correlated with the percentage of correctly perceived articulations of consonants. Furthermore, they showed unidirectional confusion patterns among consonants that were acoustically similar and reported that patients with sensorineural hearing impairment had a higher tendency to mishear consonants with a short τ_e-med for consonants with a long τ_e-med [7]. Therefore, the τ_e-med may be a useful indicator of sensorineural hearing impairment.

Table 1. Japanese monosyllables included in the 57-S monosyllable word list.

/a/	/ka/	/sa/	/ta/	/na/	/ha/	/ma/	/ya/	/ra/	/wa/	/ga/		/da/
/i/	/ki/	/si/	/ti/	/ni/	/hi/	/mi/		/ri/			/zi/	
/u/	/ku/	/su/	/ti/		/hu/	/mu/	/yu/	/ru/			/zu/	
/e/	/ke/	/se/	/te/	/ne/		/me/		/re/				/de/
/o/	/ko/	/so/	/to/	/no/	/ho/	/mo/	/yo/	/ro/		/go/		/do/

Fifty monosyllables are randomly arranged in each word list.

Since speech recognition levels vary across patients, the relationship between the percentage of accurately perceived articulation and the τ_e-med may also vary depending on speech recognition. Therefore, this relationship may not be applicable to all patients with sensorineural hearing impairment. It is necessary to determine whether this relationship is present in all individuals with sensorineural hearing impairments who have different levels of speech recognition performance. However, to the best of our knowledge, no previous study examined the influence of speech recognition performance on the relationship between the percentage of accurately perceived articulation and the τ_e-med. Hence, we assessed the influence of speech recognition performance on this relationship. In addition, similar to previous studies that examined confusion patterns among acoustically similar consonants, the present study evaluated the relationship between monosyllabic confusion and the τ_e-med and explored the influence of speech recognition performance on this relationship.

We utilized the speech audiometry results of patients who visited our hospital for hearing aid fitting to assess the relationship between the τ_e and the audibility of monosyllables for patients with general sensorineural hearing loss. These data were used as samples to determine the accuracy of perceived articulation.

2. Materials and Methods

2.1. Procedure for Determining SRS

2.1.1. Subjects

We utilized speech audiometry results of patients who visited our hospital for hearing aid fitting to assess the relationship between τ_e and audibility of monosyllables in patients with general sensorineural hearing loss. These data were used as samples to determine the accuracy of perceived articulation.

The subjects were patients who visited the Nara Medical University Hospital, Kashihara, Nara, Japan, for hearing aid fitting between September 2005 and August 2007. A total of 75 adult patients (40 women and 35 men) with sensorineural hearing loss visited the Nara Medical University Hospital in that period. Patients with sensorineural hearing loss were defined as those who had no obvious air bone gap on pure-tone audiometry.

We provided the participants with the opportunity to opt out, and this information was highlighted on the website of the Department of Otolaryngology-Head and Neck Surgery, Nara Medical University. This study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of Nara Medical University (No. 2245).

Speech audiometry was performed on a total of 75 patients. Clinically, speech audiometry is a standard test for hearing aid fitting in Japan and is performed according to the Japanese Society of Audiology's Methods of Audiometry (2003) [22]. For 6 of the 150 ears of the 75 patients, speech audiometry could not be performed, as these patients could not hear at all. For 3 of the 150 ears, speech audiometry was performed, but the patients were nearly unresponsive (the percentages of response were 0%, 0%, and 2%). Therefore, these nine ears in total were excluded, as they were not suitable for analyses. Finally, 141 ears of 75 patients were included. The mean age of the patients was 69.9 ± 13.7 years. The average hearing levels of 500, 1000, and 2000 Hz in pure-tone audiometry (PTA) was at 49.1 ± 7.2 dB.

2.1.2. Speech Audiometry Procedure

In Japan, speech audiometry is conducted for hearing aid fitting, and monosyllables, particularly 57-S, are used in the word hearing test. Speech audiometry was conducted using the 57-S and 67-S word lists (including 50 and 20 monosyllables, respectively), which are the standard lists used in Japan, authorized by the Japan Audiological Society [22,23] (Table 1).

The 57-S and 67-S word lists include five and eight monosyllable-order patterns, respectively. The monosyllable-order patterns are randomly selected. All monosyllables are spoken by a single female speaker. The speech recognition test using the 57-S word list is considered to be more informative compared with the test using the 67-S word list; moreover, the test using the 57-S word list reflects better speech recognition ability, owing to the larger amount of test materials than in the test using the 67-S word list.

However, a longer examination time is required for the repeated measurements using the 57-S word lists. To reduce the burden of measurement, speech recognition scores (SRSs) were first measured using the 67-S word list at intervals of 10 dB steps until the increase in the score was saturated. In some cases, the speech recognition curve may not reach saturation when the presentation level is increased due to possible equipment limitations in terms of output level or when an uncomfortable level of loudness for the patient is reached. In such cases, the obtained SRS is recognized as the best score but not exactly as the maximum value. The presentation level at which the maximum score was obtained in the performance-intensity function was defined as "dB (max)". After the dB (max) was identified using the 67-S word list for each participant, the SRSs at the dB (max) were measured using the 57-S word list. The obtained scores were defined as the maximum SRSs of each ear and were used for analyses.

Speech audiometry was performed using a conventional audiometer (AA-78; Rion, Tokyo, Japan) that conformed to IEC 60645-1 [24]. Speech signals were presented using earphones (AD-02T, Rion) associated with the audiometer. The earphones were calibrated using a sound-level calibrator (AG-64, Rion). The monosyllable word lists were reproduced using a CD player (CDP-XE500; SONY, Tokyo, Japan). Measurements were performed in a soundproof room.

2.1.3. Grouping

The 141 ears were classified into three groups based on the SRSs: high-SRS group, SRS 70–100%; middle-SRS group, SRS 50–68%; and low-SRS group, SRS 10–48%.

In Japan, the criterion for determining the degree of hearing impairment is determined based on a maximum SRS of less than 50% for audiometry testing. Moreover, speech audiometry under noise is not conducted in cases wherein the maximum speech recognition score is less than 50% without noise during hearing aid fitting [25]. Therefore, we considered it appropriate to categorize the subjects based on SRSs, with a threshold of 50%. Since the number of subjects in each group was uneven, we partitioned the group with a mean SRS of more than 50% into two subgroups. We obtained the percentages of accurately perceived articulations from the 50 monosyllables and of accurately perceived articulations of consonants for the three SRS groups.

2.2. ACF Factors Obtained from the Monosyllables

The ACF parameter was calculated using MATLAB R2016b (MathWorks, Natick, MA, USA) from the normalized ACF:

$$\phi(\tau) = \frac{\Phi(\tau)}{\Phi(0)}, \tag{1}$$

where

$$\Phi(\tau) = \frac{1}{2T}\int_{-T}^{T} p\prime(t)p'(t+\tau)dt, \tag{2}$$

Here, τ is the delay time (s), $2T$ is the integration interval (s), and $p'(t)$ is the signal after it is passed through an A-weighted filter. As recommended in a previous study [7], we calculated the running ACF using an integral interval of 80 ms, with 5 ms sliding steps. As an example, Figure 1 shows the logarithmic $\phi(\tau)$ value of the Japanese monosyllable /a/. τ_e is a factor of the normalized ACF, defined by a 10-percentile delay of the normalized ACF, representing repetitive features or reverberation contained within the signal itself. τ_e tracks the temporal decay of the periodicity. The represented τ_e values for each monosyllabic stimulus were determined according to τ_e-med because the obtained ACF factors were not normally distributed [7]. We also obtained τ_e-med to represent the τ_e values for each monosyllable included in the 57-S word list.

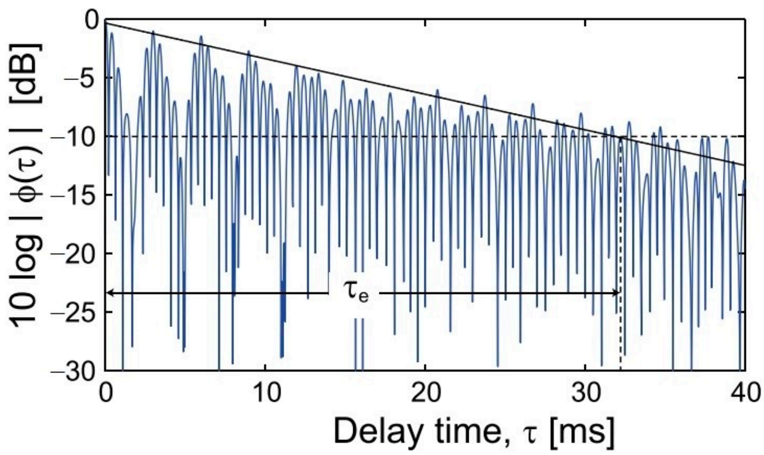

Figure 1. Logarithmic $\phi(\tau)$ value of the Japanese monosyllable /a/. The effective duration (τ_e) is the ACF factor, defined by the delay time at which the envelope along the early decay of the normalized ACF becomes −10 db.

2.3. Data Analysis

Fifty monosyllables were classified according to the consonants. The percentages of accurately perceived articulations were obtained for each consonant in the high-, middle-, and low-SRS groups. The correlations between the percentage of accurately perceived articulations and τ_e-med for each consonant in each group were analyzed using Pearson's coefficient of correlation test.

The τ_e-meds of the presented- and response-monosyllables were gathered for each patient. The average τ_e-meds of the presented- and response-monosyllables were compared among the three groups. Monosyllables with no responses were excluded from the comparison. Monosyllables were also excluded if the corresponding monosyllables were not included in the 57-S word list. Two-factor mixed analysis of variance (ANOVA) was performed to analyze the effect of group (high-, middle-, and low-SRS) and condition (presentation/response) on the averaged τ_e-med. The group was the between-participant factor, and the condition was the within-participant factor.

The percentages and τ_e-meds of "no response" monosyllables were compared among the three groups. One-factor ANOVA was performed to determine the percentage and τ_e-meds of "no response" monosyllables. Tukey's method was used for multiple comparisons. Statistical analyses were performed using IBM SPSS Statistics software version 29 (IBM, Armonk, NY, USA).

3. Results

In total, 47, 53, and 41 ears were classified into the high-, middle-, and low-SRS groups, respectively. Table 2 shows the demographic and clinical characteristics of the three SRS groups. One-way ANOVA was performed, and the results showed significant differences for the SRS factor ($F(1,138) = 317.814$, $p < 0.001$) and age ($F(1,138) = 3.230$, $p = 0.043$). Multiple comparisons using Tukey's method showed significant differences among the three groups for the SRS factor ($p < 0.001$) and between the middle- and high-SRS groups for age ($p < 0.001$).

Table 2. Patients' demographic and clinical characteristics.

	Number	Age (years)	Average PTA (dB)	SRS (%)	Presentation Level of Speech Recognition Test (dB HL)
High-SRS group	47	67.0 ± 12.5	48.6 ± 11.0	81.6 ± 7.6	84.3 ± 8.8
Middle-SRS group	53	73.3 ± 10.5	47.2 ± 5.2	59.3 ± 6.0	84.6 ± 8.7
Low-SRS group	41	71.6 ± 15.3	50.1 ± 7.0	38.3 ± 9.7	88.0 ± 9.1

High-SRS group: SRSs 70–100%; middle-SRS group: SRSs 50–68%; low-SRS group: SRSs 10–48%. Average PTA: average value of hearing levels at 500, 1000, and 2000 Hz for PTA.

Figure 2 shows the relationship between the percentage of articulation and the τ_e-med for each consonant in the high-, middle-, and low-SRS groups. Each plot in Figure 1 indicates the average percentage of articulation and the average τ_e-med of each consonant. The average τ_e-med increased as the percentage of articulation increased. Significant correlations were observed between the percentage of articulation and the average τ_e-med in the three SRS groups (high-SRS group: $rs = 0.848$, $p = 0.0001$; middle-SRS group: $rs = 0.784$, $p = 0.0009$; low-SRS group: $rs = 0.826$, $p = 0.0003$).

The averaged τ_e-meds of the presentations and responses for the three groups are shown in Figure 3. Two-factor mixed ANOVA revealed a significant main effect of the condition (presentation/response) ($F(1,138) = 83.062$, $p < 0.001$). No significant main effect of the group (high/middle/low) ($F(2,138) = 1.902$, $p = 0.153$) was observed, with no significant interaction between the groups ($F(2,138) = 1.881$, $p = 0.156$). The average τ_e-med of the response was significantly longer than that of the presentation in all three groups.

The percentage of "no response" responses differed among the three groups. The low-SRS group had a higher percentage of "no response" responses (high-SRS group: 1.0%; middle-SRS group: 1.29%; low-SRS group: 4.48%). The results showed significant differences in the SRS factor ($F(1,138) = 9.115$, $p < 0.001$). Multiple comparisons using Tukey's method showed significant differences for the following combinations (low-SRS group > high-SRS group, $p < 0.001$; low-SRS group > middle-SRS group, $p < 0.001$). The average τ_e-meds of the presentations for "no response" were 46.826, 41.937, and 47.460 in the high-, middle-, and low-SRS groups, respectively, indicating no significant difference ($F(1,292) = 3.348$, $p = 0.063$).

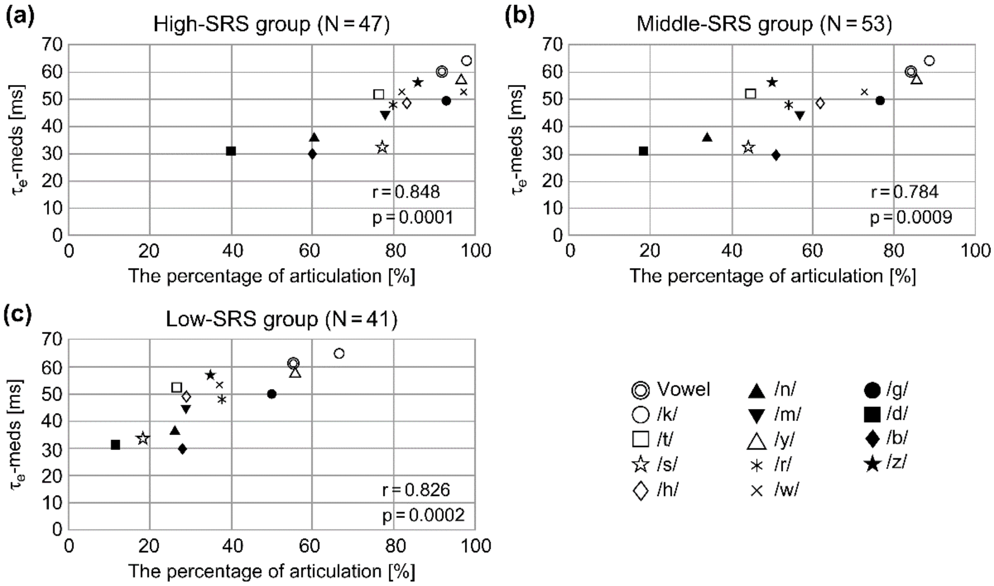

Figure 2. Relationship between the percentage of articulation and τ_e-med. Panels (**a**–**c**) show the results of high-, middle-, and low-SRS groups, respectively. The different symbols indicate different consonants. SRS, speech recognition score; τ_e-med, median effective duration.

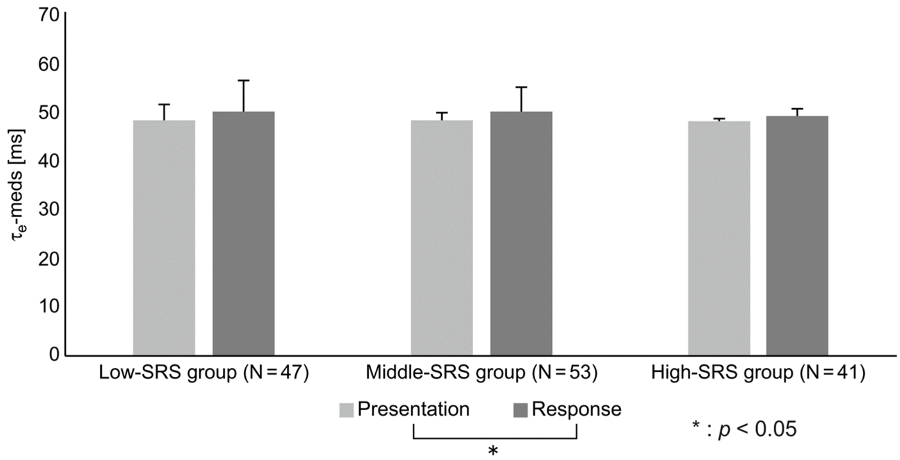

Figure 3. The averaged τ_e-meds of presentation and response. SRS: speech recognition scores; τ_e-med: median effective duration. Values are presented as means ± standard deviation of participants in high-, middle-, and low-SRS groups.

4. Discussion

4.1. Relationship between Percentage of Articulation and τ_e-med

Individuals with normal hearing can accurately identify almost all monosyllables under quiet conditions when presented at a sufficient sound pressure level. However, the accuracy of listening decreases under noise-presented conditions [26,27]. Ando et al. evaluated the effect of noise on the τ_e of monosyllables [20]. They reported that the τ_e of monosyllables decreased by up to half when measured under noisy conditions and that the τ_e was strongly correlated with the decrease in speech recognition in individuals with normal hearing. Other studies [11,19–21] showed similar results under noise and reflection conditions. The τ_e is one of the significant factors for estimating the speech recognition of individuals with normal hearing in various sound fields. Furthermore, a previous study reported that the τ_e-med is strongly correlated with the percentage of accurately perceived articulations in patients with sensorineural hearing impairment. The τ_e-med of a monosyllable represents the balance between the periodicity of the vowel and consonant components. When the vowel and consonant components of a monosyllable produce a stable wave at the fundamental frequency of the voice, the length of the τ_e-med increases [7]. Patients with sensorineural hearing impairment may recognize the stability of the temporal pattern of sound waves as a cue to identify certain monosyllables. However, previous studies did not evaluate the influence of the degree of speech recognition. Significant correlations were observed between the percentage of articulation and the average τ_e-med in the three SRS groups, suggesting that the stability and persistence of the temporal pattern of waves are important factors for speech recognition, even in patients with poor speech recognition. Hence, the τ_e-med is useful for estimating the monosyllables that patients with sensorineural hearing impairment find easy to listen to, independent of speech recognition performance.

The most significant challenge faced by patients with sensorineural hearing impairment is understanding speech in noisy conditions [25–29]. The effect of noise is more pronounced in patients with sensorineural hearing impairment than in individuals with normal hearing [26,27]. If such a relationship also holds in noisy conditions, the τ_e-med may be useful for estimating the monosyllables that patients with sensorineural hearing impairment find easy to hear in noisy environments. However, to the best of our knowledge, no previous study examined the relationship between the listening ability of such patients in noisy conditions and the τ_e-med. Further studies on the relationship between the τ_e-med and confusion in patients with sensorineural hearing impairment in noisy conditions are needed.

4.2. Average τ_e-med of the Presented Monosyllables and Responses

We compared the average τ_e-meds of monosyllables and the responses of the three groups. We found that the average τ_e-meds of the monosyllables were significantly longer than those of the responses of each of the three groups. This indicates that monosyllables tend to be misheard as monosyllables with long τ_e-meds when confusion occurs. A previous study involving patients with sensorineural hearing impairment suggested that consonant confusion patterns were unidirectional and that the τ_e-med was an influential factor in determining the direction [7]. Specifically, consonants with a short τ_e-med tend to be misheard as consonants with a long τ_e-med when confusion occurs. However, we investigated the tendency of confusion among monosyllables rather than consonants. The tendency observed in this study corroborates the results of a previous study [7]. Furthermore, this tendency does not vary depending on speech recognition performance. The results indicate that patients with clinical sensorineural hearing impairment are likely to misinterpret monosyllables with long τ_e-meds when experiencing confusion, and this tendency is not influenced by speech recognition performance. Therefore, modification of the τ_e, such as through hearing aid adjustment, would be helpful for processing speech signals. A previous study proposed a method to control the temporal fluctuations in the fundamental frequency of a voice that can be used to control the τ_e of a monosyllable [7]. When individuals with normal hearing heard a monosyllable whose τ_e was controlled longer, the stimulus was not heard as belonging to the consonant that had a long averaged τ_e-med [7]. However, no studies evaluated how patients with clinically sensorineural hearing impairment hear a monosyllable with a τ_e that was controlled. For patients who have difficulty hearing certain monosyllables, controlling the τ_e may be helpful when fitting hearing aids. Further studies are needed to verify whether patients with sensorineural hearing impairment find it easy to hear monosyllables with τ_e values that are controlled and long.

To the best of our knowledge, the relationship between the τ_e-med and "no response" was not investigated in previous studies. In this study, the percentage of "no response" was significantly higher in the low-SRS group than in the other two groups. In contrast, there was no significant difference in the τ_e-meds of presentations for "no response" among the three groups. This result suggests that "no response" is not related to the τ_e-med, although the percentage of articulation and confusion were associated with the τ_e-med. In other words, the mechanism underlying "no response" may be different from that underlying confusion. Therefore, besides the τ_e-med, factors hitherto unknown may be related to "no response".

4.3. Limitations

In the present study, only confusion between monosyllables included in the 57-word list was considered. If a patient stated that a monosyllable was not included in the 57-S word list, this item was excluded from further analysis. However, in the three groups, the percentages of monosyllables excluded from the 57-S word list because of confusion were very low: high-SRS group, 1.4%; middle-SRS group, 3.4%; low-SRS group, 2.2%. Hence, the influence of confusion on the analysis was considered minimal.

In this study, patients were grouped according to their average hearing ability. making it difficult to isolate the specific portions of hearing loss that are responsible for the inability to accurately distinguish syllables. In future studies, we aim to examine audiometry in greater detail and further explore the relationship between the frequency value and the τ_e value.

5. Conclusions

Significant correlations were observed between the percentage of articulation and the τ_e-med, independent of speech recognition performance. The τ_e-med proved to be useful for estimating the monosyllables that patients with sensorineural hearing impairment find easy to listen to, independent of speech recognition performance. Furthermore, we found that monosyllables are prone to being misheard as monosyllables with longer τ_e-med values during instances of confusion. The τ_e-med value is the influential factor for determining the direction of confusion. Further studies are needed to verify whether artificially prolonging the τ_e of monosyllables can enhance their audibility for patients with sensorineural hearing impairment.

Author Contributions: Conceptualization, S.A. and T.N.; methodology, S.A.; software, S.A. and T.N.; validation, T.N., R.S. and T.K.; formal analysis, R.S.; investigation, S.A.; data curation, T.N.; writing—original draft preparation, S.A.; writing—review and editing, S.A. and T.N.; supervision, T.K. and H.H.; project administration, T.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Ethics Committee of Nara Medical University (No. 2245). We provided the participants with the opportunity to opt out, which was highlighted on the website of the Department of Otolaryngology-Head and Neck Surgery, Nara Medical University, in 2019.

Informed Consent Statement: The requirement for patient consent was waived due to the retrospective nature of the study. Information regarding the study was made available to the public, and participants had the opportunity to opt out from the study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rosen, S. Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. B Biol. Sci.* **1992**, *336*, 367–373. [CrossRef]
2. Hou, L.; Xu, L. Role of short-time acoustic temporal fine structure cues in sentence recognition for normal-hearing listeners. *J. Acoust. Soc. Am.* **2018**, *143*, EL127–EL132. [CrossRef]
3. Moore, B.C.J.; Glasberg, B.R.; Hopkins, K. Frequency discrimination of complex tones by hearing-impaired subjects: Evidence for loss of ability to use temporal fine structure. *Hear. Res.* **2006**, *222*, 16–27. [CrossRef]
4. Bacon, S.P.; Opie, J.M. Modulation detection interference in listeners with normal and impaired hearing. *J. Speech Lang. Hear. Res.* **2002**, *45*, 392–402. [CrossRef] [PubMed]

5. Kodera, K.; Akai, S.; Hirota, E.; Miura, M.; Yabe, S. Study on consonant confusion in Japanese patients with sensorineural hearing loss. *Nippon. Jibiinkoka Gakkai Kaiho* **1993**, *96*, 1404–1409,1573. (In Japanese) [CrossRef] [PubMed]
6. Akasaka, S.; Nishimura, T.; Okayasu, T.; Hosoi, H. Percentage of correct answers to 57-S individual Japanese monosyllabic words in the hearing impaired. *Audiol. Jpn.* **2010**, *53*, 69–75. (In Japanese) [CrossRef]
7. Shimokura, R.; Akasaka, S.; Nishimura, T.; Hosoi, H.; Matsui, T. Autocorrelation factors and intelligibility of Japanese monosyllables in individuals with sensorineural hearing loss. *J. Acoust. Soc. Am.* **2017**, *141*, 1065–1073. [CrossRef]
8. Ando, Y. Subjective preferences for sound fields. In *Auditory and Visual Sensations*; Springer: New York, NY, USA, 2009; pp. 25–38.
9. Cariani, P.A.; Delgutte, B. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J. Neurophysiol.* **1996**, *76*, 1698–1716. [CrossRef]
10. Cariani, P.A.; Delgutte, B. Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *J. Neurophysiol.* **1996**, *76*, 1717–1734. [CrossRef]
11. Ando, Y. Prediction of subjective preference in concert halls. In *Concert Hall Acoustics*; Springer: Heidelberg, Germany, 1985; pp. 70–88.
12. Sato, S.; You, J.; Jeon, J.Y. Sound quality characteristics of refrigerator noise in real living environments with relation to psychoacoustical and autocorrelation function parameters. *J. Acoust. Soc. Am.* **2007**, *122*, 314–325. [CrossRef]
13. Soeta, Y.; Shimokura, R. Sound quality evaluation of air-conditioner noise based on factors of the autocorrelation function. *Appl. Acoust.* **2017**, *124*, 11–19. [CrossRef]
14. Kitamura, T.; Shimokura, R.; Sato, S.; Ando, Y. Measurement of temporal and spatial factors of a flushing toilet noise in a downstairs bedroom. *J. Temporal Des. Archit. Environ.* **2002**, *2*, 13–19.
15. Fujii, K.; Soeta, Y.; Ando, Y. Acoustical properties of aircraft noise measured by temporal and spatial factors. *J. Sound Vib.* **2001**, *241*, 69–78. [CrossRef]
16. Fujii, K.; Atagi, J.; Ando, Y. Temporal and spatial factors of traffic noise and its annoyance. *J. Temporal Des. Archit. Environ.* **2002**, *2*, 33–41.
17. Soeta, Y.; Shimokura, R. Survey of interior noise characteristics in various types of trains. *Appl. Acoust.* **2013**, *74*, 1160–1166. [CrossRef]
18. Jeon, J.Y.; Sato, S. Annoyance caused by heavyweight floor impact sounds in relation to the autocorrelation function and sound quality metrics. *J. Sound Vib.* **2008**, *311*, 767–785. [CrossRef]
19. Korenaga, Y.; Ando, Y. A method of calculating intelligibility of sound field in relation to temporal structure of reflections—On the trend of syllable confusion under sound field composed of a direct sound and up to two reflections. *J. Acoust. Soc. Jpn.* **1996**, *52*, 940–947. (In Japanese)
20. Ando, Y. [Application]; Vol. II. Speech reception in sound fields. In *Auditory Vis Sensations*; Springer: New York, NY, USA, 2009; pp. 179–197.
21. Ando, Y.; Sakai, H.; Sato, S. Formulae describing subjective attributes for sound fields based on a model of the auditory-brain system. *J. Sound Vib.* **2000**, *232*, 101–127. [CrossRef]
22. Japan Audiological Society. Methods of speech audiometry. *Audiol. Jpn.* **2003**, *46*, 621–637. (In Japanese) [CrossRef]
23. Kodera, K.; Hosoi, H.; Okamoto, M.; Manabe, T.; Kanda, Y.; Shiraishi, K.; Sugiuchi, T.; Suzuki, K.; Tauchi, H.; Nishimura, T.; et al. Guidelines for the evaluation of hearing aid fitting (2010). *Auris Nasus Larynx* **2016**, *43*, 217–228. [CrossRef]
24. *IEC 60645-1*; Electroacoustics—Audiometric Equipment—Part 1: Equipment for Pure-Tone and Speech Audiometry. IEC: Geneva, Switzerland, 2001. Available online: https://webstore.iec.ch/publication/32370 (accessed on 10 May 2023).
25. Meister, H.; Lausberg, I.; Kiessling, J.; Walger, M.; von Wedel, H. Determining the importance of fundamental hearing aid attributes. *Otol. Neurotol.* **2002**, *23*, 457–462. [CrossRef] [PubMed]
26. Prosser, S.; Turrini, M.; Arslan, E. Effects of different noises on speech discrimination by the elderly. *Acta Oto-Laryngol.* **1990**, *476*, 136–142. [CrossRef] [PubMed]
27. Summers, V.; Makashay, M.J.; Theodoroff, S.M.; Leek, M.R. Suprathreshold auditory processing and speech perception in noise: Hearing-impaired and normal-hearing listeners. *J. Am. Acad. Audiol.* **2013**, *24*, 274–292. [CrossRef] [PubMed]
28. Kortlang, S.; Mauermann, M.; Ewert, S.D. Suprathreshold auditory processing deficits in noise: Effects of hearing loss and age. *Hear. Res.* **2016**, *331*, 27–40. [CrossRef]
29. Gatehouse, S.; Noble, W. The speech, spatial and qualities of hearing scale (SSQ). *Int. J. Audiol.* **2004**, *43*, 85–99. [CrossRef]