Research Article

Effects of Age at Onset and Years of Musical Training on Consonant Recognition in Musicians

Ebtesam Sajjadi[1,2], Ahmadreza Nazeri[1], Ali Mohammadzadeh[1], Nooshin Sayadi[3*], Saeid Farahani[3], Ghassem Mohammadkhani[3]

[1.] Department of Audiology, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[2.] Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland USA

[3.] Department of Audiology, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran

ORCID ID:
Ebtesam Sajjadi: 0000-0002-2285-503X
Ahmadreza Nazeri: 0000-0001-7234-5479
Ali Mohammadzadeh: 0000-0003-2799-4562
Nooshin Sayadi: 0000-0002-4785-3911
Saeid Farahani: 0000-0003-2541-283X
Ghassem Mohammadkhani: 0000-0002-0985-0779

Citation: Sajjadi E, Nazeri A, Mohammadzadeh A, Sayadi N, Farahani S, Mohammadkhani G. Effects of Age at Onset and Years of Musical Training on Consonant Recognition in Musicians. Aud Vestib Res. 2025;34(1):?-?.

Article info:
Received: 06 Jan 2024
Revised: 22 Apr 2024
Accepted: 18 May 2024

[*] **Corresponding Author:** Department of Audiology, School of Rehabilitation, Tehran University of Medical Sciences, Tehran, Iran. n-sayadi@razi.tums.ac.ir

Short running title: Effects of Age at Onset and Years of Musical…

Highlights:

- Starting musical training at an early age can improve consonant recognition in noise
- The years of training cannot predict the output of consonant recognition in noise
- Early musical training is crucial for improving speech perception abilities

ABSTRACT

Background and Aim: Previous research has shown that musicians have better performance in consonant recognition in noise compared to non-musicians. This study aimed to determine the effects of age at onset and years of musical training on consonant recognition in noise in musicians.

Methods: Thirty-six Persian-speaking young musicians with normal hearing whose age at onset of musical training ranged 4–8 years and had at least 10 years of musical training, were asked to listen to Consonant-Vowel-Consonant (CVC) tokens presented in 12-talker babble noise and transcribe them. Multiple linear regression

analysis was used to determine the ability of age at onset of musical training and years of musical training in predicting the recognition of stops and fricatives.
Results: The age at onset of musical training significantly predicted consonant recognition in noise ($p<0.001$, $R^2= 0.614$), while years of musical activity had no significant effect ($p=0.055$).
Conclusion: Early start of musical training seems to improve the ability to recognize consonants in challenging listening environments.
Keywords: Speech in noise; speech perception; musical training; age of acquisition; consonant recognition

Introduction

Studies have shown structural and functional brain differences by starting musical training at early ages or years of consistent musical training [1, 2]. Early musical training affects the development of the brain structure, especially in the subcortical and cortical regions [3, 4]. Musicians who begin musical training at an early age (having more experience) are more likely to have robust auditory brainstem function. Auditory Brainstem Response (ABR) continues to change throughout childhood with the developmental inflection point (where the curvature of the trajectory changes signs) usually occurring at ages 5–11. After this inflection point, the developmental trajectory returns back to the adult value and then stabilizes. It has been shown that musical training can intensify neural activity during sensitive phases of auditory brainstem development, resulting in improvements in certain elements of the complex ABR, mostly in the musicians aged 2–5 years whose response consistency is higher than in those aged 5–8 years from the general population [5].
Precise sound processing by musicians in both music and speech domains can result in better encoding of speech sounds (i.e. consonants) in quiet and in noise [6, 7]. Electrophysiological studies have proposed a neural foundation for the advantage observed in musicians, employing measures such as speech ABR and the frequency-following response. These studies have revealed a high subcortical speech discrimination among musicians, with this advantage being correlated with their Speech-in-Noise (SIN) perception [8, 9]. The advantage observed in musicians regarding SIN perception has been reported by both behavioral and electrophysiological studies. Behavioral studies have evaluated SIN perception using three common tests (words-in-noise test, Quick-Speech in-Noise test [Quick-SIN], and the Hearing-in-Noise Test [HINT]). Participants in these tests are asked to detect target words or sentences when there are informational or energetic maskers. The majority of these studies have indicated that musicians exhibit superior performance in the SIN perception [9, 10]. Accurate perception of consonants forms the foundation of overall speech intelligibility and provides crucial acoustic cues for word comprehension, particularly in challenging environments. The duration and the Voice Onset Time (VOT) are two important auditory cues for consonant recognition [11, 12]. Changes in VOT enable the distinction between voiced (e.g. /b/) and voiceless (e.g. /p/) stop consonants. The VOT is defined as the time between the release of a stop consonant and the onset of vocal cord vibration [12]. There is evidence that the endurance of structural and functional alterations induced by musical training across the auditory system not only enhance acoustic precision in response to cues such as duration, but also refine various aspects of speech processing, including VOT [9]. Musical training in various forms can also enhance sensorimotor, emotional, cognitive, and social functions [13, 14].
Several studies have indicated the effects of early musical training and years of consistent musical practice on the improvement of SIN perception in musicians. Transfer of training between music and speech depends more on common processing, attention, and memory, rather than the type of musical instrument [6]. Ruggles et al. [10] and Parbery-Clark et al. [15] showed a positive correlation between years of consistent musical training and clinical measures of SIN perception (Quick-SIN, HINT, or both). Zendel et al. [11] demonstrated that early-onset musicians perceived more words correctly in noise. Coffey et al. [16] also indicated a positive association between the age at onset of musical training and HINT scores. Precise perception of consonants provides essential acoustic information that is substantial to understanding the meaning of words, especially under difficult and noisy conditions [17, 18].
This study aimed to explore the effects of age at onset of musical training and years of consistent musical experience on consonant recognition in noise. Persian stops and fricatives were used, which constituted a maximum number of Persian consonants, embedded in twelve-talker babble noise. It is hypothesized that formal musical training during early childhood with continual practice over years can positively affect the perception of words and syllables (ie., consonants) in the presence of background noise.

Methods

Participants

In this study, participants were 36 right-handed Persian-speaking musicians aged 20–30 years (Mean age: 25.16±3.20 years), with onset of musical training at ages 4–8 years, and 12–25 years of musical experience. They were recruited using criteria based on the previous study [19]. Participants had normal hearing (pure tone threshold <25 dB for octaves at 250–8000 Hz frequency), normal speech recognition thresholds (<25 dB HL), normal middle ear function (An-type tympanometry and middle ear pressure ranging from +50 to –100 daPa), with presence of ipsilateral and contralateral acoustic reflexes (80–100 dB at 500–3000 Hz frequency), and no history of neurological, audiological, speech or linguistic disorders. The inclusion criteria were: Having at least 10 years of formal musical training (trained by attending public or private music courses for one or more instruments) [20, 21], playing an instrument before the age of 9 [5], and musical practicing for at least three times per week for three years before participation in this study [20].

Materials

All auditory assessments were performed by a two-channel audiometer (Astra, GN Otometrics, Denmark), a TDH39 headphone, and a tympanometer (AT235, Interacoustics, Denmark). Phonetic and auditory speech discrimination tests (speech reception threshold and speech discrimination score) were performed to ensure the correct speech production and normal speech discrimination ability of the participants. The stimuli consisted of 128 target monosyllabic Persian words presented concurrently with 12-talker babble noise, developed by Ahmadi et al. [22], at a Signal-to-Noise Ratio (SNR) of –5 dBHL. When babble noise consists of one or two talkers, listeners with normal hearing can utilize gaps in the masking noise to identify the target speaker. However, as the number of competing talkers increases (e.g., an increase to 12 talkers in our study), masking also increases and the gaps in the temporal envelope decrease. Nonetheless, listeners may still be able to utilize time-varying cues of multi-talker babble to distinguish the target talkers [23, 24].

All words were selected from the lists developed by Mosleh [25], which were meaningful monosyllabic words in Consonant-Vowel-Consonant (CVC) formant (e.g. miz (“chair”), sib (“apple”) etc.). These words were phonetically balanced and included stops (/b/, /p/, /t/, /d/, /g/, /k/, /q/, /ʔ/) and fricatives (/s/, /z/, /š/, /x/, /ž/, /v/, /f/, /h/). Attempts were made to include all possible CVCs composed of 16 target consonants at initial and final positions, combined with 6 vowels (/i/, /â/, /e/, /u/, /o/, /a/). Some of the CVCs (26 out of 128) were considered for both initial and final positions, while the rest were only considered for one position. A professional male speaker with expertise in phonetic science recorded these stimuli in a professional recording studio. Adobe Audition CS5.5 v4.0 software was used to perform the experiment and the stimuli were presented directly by an audiometer (Astra, GN Otometrics, Denmark). To facilitate handwritten responses and increase trial precision, a four-second interval between words was considered.

Procedure

Participants completed all tasks in a single one-hour session, supervised by the audiologist. Prior to the trial, musicians responded to questions about risk factors for hearing health, use of ototoxic medications, neurological disorders, and history of smoking. Otoscopy was then conducted to ensure normal external ear conditions and being free from earwax or debris. Tympanometry and both ipsilateral and contralateral acoustic reflexes were measured. Pure tone and speech recognition thresholds were individually assessed in a quiet room with minimum ambient noise. Before the main trial, participants underwent phonetic and auditory discrimination tests. The right ear was designated as the Test Ear (TE), with both noise and CVCs presented ipsilaterally. Participants were instructed to ignore the noise and transcribe the keywords they identified from the background multi-talker babble noise. Stimuli were presented at 30 dB above the speech recognition threshold (at a comfortable hearing level). Participants first listened to monosyllabic words in quiet and then to twelve-talker babble noise at a –5 SNR. Responses were considered correct only if the listener successfully identified the corresponding target consonant followed by the target vowel in the initial or final position from the original list. Finally, the percentage of correctly recognized stops and fricatives was calculated separately for each participant [20].

Data analysis

The relationship between musical training variables (age at onset of musical training and years of training) and stop and fricative recognition scores were examined using the Pearson correlation test. To predict the effects of age at onset of musical training and years of musical training on the recognition of stops and, multiple linear regression analysis was performed and the adjusted R^2 values were reported. Shapiro-Wilk test and assessment of kurtosis and skewness were used to ensure the normality. All statistical analyses were performed in IBM SPSS v.17 software.

Results

A total of 36 musicians (20 male and 16 female) aged 20–30 years participated in this study. Descriptive statistics (mean, standard deviation, minimum and maximum) for variables in current study, are shown in Table 1. Results of the correlational test revealed a significant correlation between musical training variables and stop and fricative recognition scores (Figure 1).

Two regression models were used to assess the ability of age at onset of musical training and years of musical training to predict the scores of stops and fricatives recognition. A significant regression correlation was found for the recognition of fricatives ($F_{(2,33)}$=26.28, p<0.001, R^2=0.614). In this model, only the age at onset of musical activity was statistically significant; the years of training was not found to be significant in this model (p=0.055). The fricatives recognition scores were equal to 40.97–3.52 (for the age at onset). The indicates that the score of fricatives recognition increased by 3.52% for every one-year decrease in the age at onset of musical training.

For the recognition of stop consonants, a significant regression correlation was also reported ($F_{(2,33)}$=16.14, p<0.001, R^2=0.495). Similar to the previous model, only the age at onset of musical training was statistically significant in this model, while years of training was not significant (p=0.09). The stops recognition scores were equal to 32.79–2.64 (for the age at onset). This indicates that the score of stops recognition increased by 2.64% for every one-year decrease in the age at onset of musical training.

Discussion

The purpose of the present study was to examine the relationship between consonant recognition in noise and two influential factors of age at onset of musical training and years of musical training in young Iranian musicians. A very strong association was found between age at onset of musical training and consonant recognition. This result supports the results of previous studies where it was shown that SIN perception was positively affected by early start of musical training [11, 16]. A possible explanation for this finding is that children acquire consonants during the early years of life. Five-year-old children have acquired most consonants within their native language [26]. Auditory experience during these years contributes to a biological basis for sound processing, and has a dramatic impact on neural structure and function. In addition, the acoustic complexity of music makes it a robust tool for influencing neural plasticity during the first years of auditory development; therefore, musical training during this period has the maximum capacity to shape the basic features of auditory function [27, 28].

The lateral surface of left superior temporal gyrus, the frontoparietal network, and the bilateral parietal cortex are involved in consonant recognition. Evidence has shown that starting musical practice at an early age can lead to greater enlargement of these regions [29, 30], and may enhance their function in processing of consonants.

Musicians, compared to non-musicians, are more sensitive to acoustic cues that are common in both music and speech domains (i.e. duration, frequency, intensity, and timbre). It has been shown that an early start of musical training not only can improves the processing of acoustic parameters such as duration, frequency and intensity by the auditory nervous system, but also can facilitate the perception of changes in the VOT [31]. Based on the results, the age at onset of musical training can be a good predictor of consonant recognition in noise.

Consistent with previous studies, we did not find a significant association between years of musical training and consonant recognition in noise [3, 10, 15]. Some other studies also failed to obtain a significant association between years of musical training and SIN perception [11]. There are certain methodological factors that may contribute to this finding. In our study, at least three times of musical training per week was considered as an inclusion criterion. The intensity of training may be was different among musicians, which affected their performance. A previous study found that the hours of musical training before the age of 11 was positively associated with neural plasticity and superior SIN perception in musicians [32]. Thus, musicians with fewer years of musical training may instead benefited from extensive hours of practice during an optimal age range. The factors such as family support and constant exposure to music at home are also effective [33], which can positively affect the SIN perception in musicians with fewer years of musical training. Furthermore, in our study, the age at onset of musical training in musicians varied from 4 to 8 years. Considering the significant effect of the early start of musical training on consonant recognition, it is suggested that the age at onset should be controlled to conduct a more precise study on the effect of years of musical training on consonant recognition in future research.

Due to the absence of clinical cognitive evaluation in this study, it was not possible to assess the association between consonant recognition and cognitive function. It is recommended that future studies include cognitive function assessments. Moreover, this study involved musicians proficient in some classical and Persian musical instruments; however, we did not assess the impact of the instrument type on consonant recognition. Conducting similar studies on individuals experienced in playing other instruments are recommended. Future studies are also recommended to include and compare two groups of musicians with early and late music learning to confirm the effects of early musical training on speech perception.

Conclusion

Musical training during optimal age of auditory development can be beneficial for consonant recognition in noise. Therefore, early musical training is recommended for people with deficits in the processing of auditory information and for children with developmental language disorders such as dyslexia, who have difficulty in consonant recognition. Although there is a positive relationship between years of musical training and consonant recognition in noise, the years of musical training cannot predict consonant recognition in noise.

Ethical Considerations

Compliance with ethical guidelines

This study was approved by Shahid Beheshti University of Medical Sciences (No: IR. SBMU. RETECH. REC. 1396. 569). All participants gave their informed consent prior to the administration of the study.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors' contributions

ES: Study design, data collection and conducting experiments, writing the original manuscript; AN: Reviewing and finalizing the manuscript, supervised the study; AM: Reviewing and finalizing the manuscript, supervised the study; NS: Data analysis and interpretation of results, writing the original manuscript; SF: Contribution to the design and drafting of suggestions for future studies; GM: Reviewing and finalizing the manuscript.

Conflict of interest

The authors have no conflict of interest to declare.

Acknowledgments

The authors would like to thank all participants for their contributions to this work.

References

1. Bermudez P, Lerch JP, Evans AC, Zatorre RJ. Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. Cereb Cortex. 2009;19(7):1583-96. [DOI:10.1093/cercor/bhn196]
2. Bengtsson SL, Nagy Z, Skare S, Forsman L, Forssberg H, Ullén F. Extensive piano practicing has regionally specific effects on white matter development. Nat Neurosci. 2005;8(9):1148-50. [DOI:10.1038/nn1516]
3. Parbery-Clark A, Skoe E, Kraus N. Musical experience limits the degradative effects of background noise on the neural processing of sound. J Neurosci. 2009;29(45):14100-7. [DOI:10.1523/JNEUROSCI.3256-09.2009]
4. Shahin AJ. Neurophysiological influence of musical training on speech perception. Front Psychol. 2011;2:126. [DOI:10.3389/fpsyg.2011.00126]
5. Skoe E, Kraus N. Musical training heightens auditory brainstem function during sensitive periods in development. Front Psychol. 2013;4:622. [DOI:10.3389/fpsyg.2013.00622]
6. Besson M, Chobert J, Marie C. Transfer of Training between Music and Speech: Common Processing, Attention, and Memory. Front Psychol. 2011;2:94. [DOI:10.3389/fpsyg.2011.00094]
7. Musso M, Fürniss H, Glauche V, Urbach H, Weiller C, Rijntjes M. Musicians use speech-specific areas when processing tones: The key to their superior linguistic competence? Behav Brain Res. 2020;390:112662. [DOI:10.1016/j.bbr.2020.112662]
8. Coffey EB, Herholz SC, Chepesiuk AM, Baillet S, Zatorre RJ. Cortical contributions to the auditory frequency-following response revealed by MEG. Nat Commun. 2016;7:11070. [DOI:10.1038/ncomms11070]

9. Parbery-Clark A, Tierney A, Strait DL, Kraus N. Musicians have fine-tuned neural distinction of speech syllables. Neuroscience. 2012;219:111-9. [DOI:10.1016/j.neuroscience.2012.05.042]
10. Ruggles DR, Freyman RL, Oxenham AJ. Influence of musical training on understanding voiced and whispered speech in noise. PLoS One. 2014;9(1):e86980. [DOI:10.1371/journal.pone.0086980]
11. Zendel BR, Tremblay CD, Belleville S, Peretz I. The impact of musicianship on the cortical mechanisms related to separating speech from background noise. J Cogn Neurosci. 2015;27(5):1044-59. [DOI:10.1162/jocn_a_00758]
12. Lisker L, Abramson AS. Some effects of context on voice onset time in English stops. Lang Speech. 1967;10(1):1-28. [DOI:10.1177/002383096701000101]
13. Musacchia G, Sams M, Skoe E, Kraus N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci U S A. 2007;104(40):15894-8.
14. Lee KM, Skoe E, Kraus N, Ashley R. Selective subcortical enhancement of musical intervals in musicians. J Neurosci. 2009;29(18):5832-40. [DOI:10.1523/JNEUROSCI.6133-08.2009]
15. Parbery-Clark A, Skoe E, Lam C, Kraus N. Musician enhancement for speech-in-noise. Ear Hear. 2009;30(6):653-61. [DOI:10.1097/AUD.0b013e3181b412e9]
16. Coffey EBJ, Chepesiuk AMP, Herholz SC, Baillet S, Zatorre RJ. Neural Correlates of Early Sound Encoding and their Relationship to Speech-in-Noise Perception. Front Neurosci. 2017;11:479.
17. Li N, Loizou PC. The contribution of obstruent consonants and acoustic landmarks to speech recognition in noise. J Acoust Soc Am. 2008;124(6):3947. [DOI:10.1121/1.2997435]
18. Kewley-Port D, Burkle TZ, Lee JH. Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. J Acoust Soc Am. 2007;122(4):2365-75. [DOI:10.1121/1.2773986]
19. Boebinger D, Evans S, Rosen S, Lima CF, Manly T, Scott SK. Musicians and non-musicians are equally adept at perceiving masked speech. J Acoust Soc Am. 2015;137(1):378-87. [DOI:10.1121/1.4904537]
20. Sajjadi E, Mohammadzadeh A, Sayadi N, Nazeri A, Tabatabai SM. Comparison of the recognition scores of stop and fricative consonants in babble noise between musicians and non-musicians. Psychol Music. 2021;49(5):1344-55. [DOI:1177/0305735620953616]
21. Schön D, Tillmann B. Short- and long-term rhythmic interventions: perspectives for language rehabilitation. Ann N Y Acad Sci. 2015;1337:32-9. [DOI:10.1111/nyas.12635]
22. Ahmadi A, Fatahi J, Keshani A, Jalilvand H, Modarresi Y, Jalaie S. [Developing and evaluating the reliability of acceptable noise level test in Persian language]. J Rehab Med. 2015;4(2):109-17. Persian.
23. Buus S. Release from masking caused by envelope fluctuations. J Acoust Soc Am. 1985;78(6):1958-65. [DOI:10.1121/1.392652]
24. Carhart R, Tillman TW, Greetis ES. Perceptual masking in multiple sound backgrounds. J Acoust Soc Am. 1969;45(3):694-703. [DOI:10.1121/1.1911445]
25. Mosleh M. [Development and Evaluation of a Speech Recognition Test for Persian Speaking Adults]. Audiol. 2001;9(1-2):72-6. Persian.
26. McLeod S, Crowe K. Children's Consonant Acquisition in 27 Languages: A Cross-Linguistic Review. Am J Speech Lang Pathol. 2018;27(4):1546-71. [DOI:10.1044/2018_AJSLP-17-0100]
27. Strait DL, Kraus N, Skoe E, Ashley R. Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. Eur J Neurosci. 2009;29(3):661-8. [DOI:10.1111/j.1460-9568.2009.06617.x]
28. Strait DL, Parbery-Clark A, O'Connell S, Kraus N. Biological impact of preschool music classes on processing speech in noise. Dev Cogn Neurosci. 2013;6:51-60. [DOI:10.1016/j.dcn.2013.06.003]
29. Carreiras M, Price CJ. Brain activation for consonants and vowels. Cereb Cortex. 2008;18(7):1727-35. [DOI:10.1093/cercor/bhm202]
30. Vaquero L, Hartmann K, Ripollés P, Rojo N, Sierpowska J, François C, et al. Structural neuroplasticity in expert pianists depends on the age of musical training onset. Neuroimage. 2016;126:106-19. [DOI:10.1016/j.neuroimage.2015.11.008]
31. Chobert J, Marie C, François C, Schön D, Besson M. Enhanced passive and active processing of syllables in musician children. J Cogn Neurosci. 2011;23(12):3874-87. [DOI:10.1162/jocn_a_00088]
32. Cho E. Sensitive periods for music training from a cognitive neuroscience perspective: A review of the literature with implications for teaching practice. International Journal of Music in Early Childhood. 2019;14(1):17-33. [DOI:10.1386/ijmec.14.1.17_1]
33. Penhune VB. Sensitive periods in human development: evidence from musical training. Cortex. 2011;47(9):1126-37. [DOI:10.1016/j.cortex.2011.05.010]

Table 1. Descriptive data for stop and fricative recognition scores and musical training variables (N=36).

Variable	Mean	SD	Minimum	Maximum
Stop recognition score*	24.00	6.17	12.50	39.20
Fricative recognition score*	28.00	7.19	15.60	46.40
Years of training	18.88	3.37	12.00	25.00
Age of musical training onset (year)	6.27	1.38	4.00	8.00

* Measured in percent

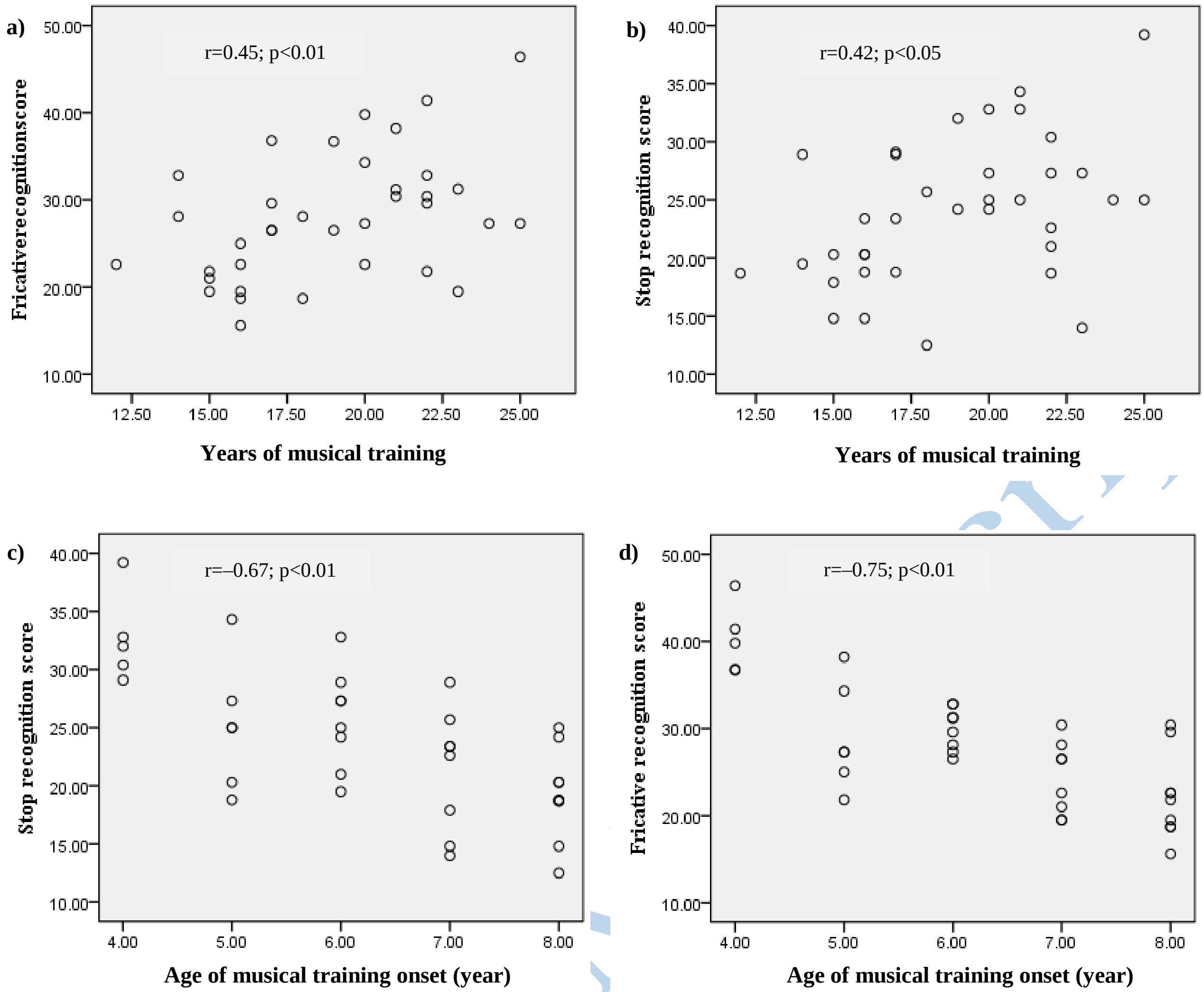

Figure 1. Correlation between age of musical training and a) stop and b) fricative recognition scores, and between years of musical training and c) stop and d) fricative recognition scores