ABSTRACT

A research project was conducted with the aim to examine whether music education may be viewed as one of the factors which improve second language acquisition. Two groups of subjects: professional musicians and non-musicians were recruited and their responses to foreign language stimuli were recorded and then examined.

1. INTRODUCTION

A number of studies revealed that a range of factors affects second language acquisition and various processes take place during the acquisition. Although the first component of language development, which is appropriate brain and the whole nervous system organization, seems to be crucial, however, several other factors such as e.g. environmental, emotional and motivational ones cannot be omitted.

Only a limited number of studies examined a possible impact of music education on language acquisition (e.g. [1], [3], [6], [7]). Moreover, there is a still ongoing discussion on the level of relationship. Namely, it is examined whether music education, music exposure or musicality improve human potential in language acquisition [4].

When looking for the possible transfer between musical training and language several approaches have been proposed. The approaches are mainly based on the fact that training in music requires engagement and refinement of processes involved in the analysis of pitch patterns over time and then the processes may be activated during interpretation of emotions conveyed by spoken utterances (cf. [1], [6], [7]). Indeed, some recent studies have provided evidence confirming the relationship. Some of the processes are shared by both domains (e.g. discrimination of emotional meaning, acoustical cues), several of them are domain-specific. To date the issue has been noticed in several studies (e.g. by Thompson, Schellenberg and Husain [6], [7]).

For instance, in two of their experiments Thompson, Schellenberg and Husain [6] examined the hypothesis that music lessons generate positive transfer effects that influence speech perception. The authors provided evidence that musically trained participants outperformed untrained participants in extracting prosodic information from speech and they suggested the existence of cognitive transfer between music and speech. They have also claimed that music lessons improve the ability to extract prosodic cues as well as the ability to interpret speech prosody.

2. RESEARCH DESIGN

A research project was conducted with the aim to investigate relationship between music education and second language acquisition. The focus was given to sounds and construct perception and production. The main goal of the study was to examine whether active involvement in music has influenced second language acquisition.

2.1. The corpus

82 word sequences in 6 languages: American English (15), British English (14), Belgian Dutch (11), French (10), Italian (10), European Spanish European (6), South American Spanish (4), and Japanese (10) have been synthesized for the corpus. The ScanSoft® RealSpeak™ application was used for the purpose.

Languages were chosen according to the typological classification. Stimuli involved both stress-timed, syllable-timed and moraes-timed languages. Amongst the sentences were questions, statements, and orders. The corpus also contained some phonological words, names and/or other short word sequences.

All word sequences were recorded on CD (3 times repeated with short gaps left between the repetitions, and a longer pause inserted in order to provide speakers with time needed to repeat the sentence). The corpus was used in further data collection.

2.2. Participants

A group of 106 subjects, Polish native speakers of Polish, with and without musical education and training and with and without different languages competence, was examined. All subjects were recruited in Lodz and Kutno areas and participated in the study on a voluntary basis. Subjects were aged from 15 to 69 years (mean 32, median 28).

2.3. Questionnaire

For the purpose of the study a special questionnaire was developed. The questionnaire included information on
participants’ sex, age, education (including start date of musical education and training as well as contact with foreign languages), music exposure, occupation, job, interests, and health (subjects were asked to give information on previous hearing problems and all illnesses which could negatively affect hearing).

2.4. Main procedure

Subjects’ ability to imitate foreign language phrases was tested. The task was meant to examine an ability to integrate different components of linguistic information such as: phonology, syntax, and intonation. The job was not a pure measure of the enumerated components but it was rather aimed at finding a key to success or failure in the language sounds and structure acquisition (perception and production).

Subjects were asked to repeat as accurately as they could some synthetic foreign language word sequences played on a CD player (Grundig) placed in a quiet area. No other information was given to the subjects. Examinees were not informed that they heard synthetic stimuli. Subjects’ productions were recorded with Sharp MD-MT200 portable recorder and UNITRA-Tonsil Microphone MCU-53 with a linear characteristic.

All recordings were carefully listened to and analyzed. The main goal was to determine whether subjects with different musical expertise perform at the same or similar level.

2.5. Research results

The author rated the speech samples by auditory analysis. Recordings were examined in a randomized order and after a period of more than one year from data collection so that to ensure unbiased evaluation of all performances.

In the first round of data analysis the scoring procedure was based mainly on a general review and observation whether all speakers responded to the stimuli and were able to repeat the speech material in the given time and accurately. It was noticed that almost all subjects encountered difficulty with at least one sentence.

In order to evaluate whether the task was not too difficult the Difficulty Factor, which optimal level equals 0.5 and which is usually used to check the proportion of respondents who were able to give the right answer to a given question or task, was calculated.

The difficulty factor may be calculated using the following formula:

\[ D = \frac{c}{n} \]  

D - difficulty factor,

c - number of correct answers,

n - number of respondents

As the main purpose of the study was to discriminate between different levels of performance, thus items with difficulty values between 0.3 and 0.7 would be most effective. In the study, the factor shows that the applied procedure and its difficulty were close to optimum and the task was feasible. Namely the factor equals 0.56, in case of musicians, and 0.39, in non-musicians which means that the task was available for both groups of speakers.

Not all subjects were able to repeat all the stimuli. The mean number of correct repetitions (i.e. these very close to the original samples) was 45.95. It should be noted that data presented in the paper refer to the stimuli taken as whole word sequences. It means that even a very slight error caused to admit a production to be incorrect.

As showed in Figure 1, musicians encountered fewer difficulties in speech repetition and produced 56.53 of correct responses to 82 provided stimuli.

![Correct Responses](image)

**Figure 1.** Mean number of correct responses.

Non-musicians performed significantly worse than musicians and produced 39.91 of correct repetitions. It means that 65.53% of musicians’ and 46.55% of non-musicians’ productions were rated as correct.

In Figure 2 below, the graph with all correct performances of all participants of the study is presented.

![Distribution of Correct Performances](image)

**Figure 2.** Number of correct performances.

The presented data may suggest that musicians could have better memory, and this parameter enabled them to perform better during the whole study. They just encountered fewer difficulties with remembering speech passages thus it may be assumed that they encountered fewer boundaries with the task.
It was found that a number of correct productions differed among languages. It was reported that most musicians repeated all stimuli on time, however not all productions were fully faithful to the original.

Detailed analysis of all questionnaires revealed that 14 non-musicians had had in the past some musical background. Therefore, all subjects were divided into four groups: without any musical training in the past, above 0 to 6 years of music education, from 7 to 12 years of musical expertise and more than 12 years.

In Figure 3 are presented scores obtained by participants of the study grouped in accordance with the length of musical training. The graphs reveal that even several years of musical education in the past affected the level of performance in the study. This result clearly shows that music training influences humans’ ability to perceive and produce foreign language speech sequences.

![Figure 3. Scores obtained by subjects with different musical background.](image)

### 2.6. More detailed data analysis

As one of the aims of the study was to establish what types of errors were produced by participants, the recorded data were listened to and as far as possible all inconsistencies and errors produced by the speakers were analyzed and assessed. Special attention was given to all mispronunciations that occurred systematically and in several speakers in the same words’ or sounds’ sequences.

Some mispronunciations and inconsistencies were observed at both segmental and suprasegmental levels (cf [5]). It was determined that many subjects, more likely non-musicians than musicians, changed several segments, repeated word sequences closer to Polish pronunciation, and did not follow appropriate production in foreign languages.

Moreover, it was observed that the modifications referred to vowels (e.g. their quality and length), consonants, and consonantal clusters as well. Interestingly, the least problems were encountered for intonation as both groups performed at similar level.

Several productions could be described as completely unintelligible in term of segmental level, however, with appropriate mimicry of speech melody. This may resulted of the so called phonemic restoration phenomenon observed by Kashino [2], among others. Namely, he claimed that “the sounds we hear are not copies of physical sounds” and “what we perceive is the result of [an] unconscious interpretation” [2].

Most commonly present mispronunciations that could be observed in almost all the word sequences were: lack of differentiation of the length of vowels occurring in a given sentence, change of vowels’ quality, difficulties with repetition of longer or more complex sentences, replacement of voiced consonants into voiceless ones and vice versa, respectively. In many cases subjects were not able to repeat either whole words or their parts (e.g. syllables, segments).

Errors of segmentation were an important source of mistakes. It was observed that mispronunciation of one segment (e.g. a consonant) resulted in other mispronunciations in neighboring segments (e.g. vowel), and vice versa. It should be also pointed out that different types of errors occurred in almost all participants’ productions and almost all of them encountered some difficulties in appropriate realization of the task as a whole.

Subjects produced both errors of performance (slips of the tongue) and errors of competence (pronunciation) [5]. Moreover, in subjects’ mimicry occurred some both native and foreign interferences.

### 2.7. Correlations and statistical analysis

Pearson correlations were performed in order to establish relationship between results achieved by speakers and their musicianship, results achieved in the memory test for music stimuli, number of years of music education, number of years of learning foreign languages, and results of the whole test of music abilities.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Variable</th>
<th>Correlation vs. probability level</th>
</tr>
</thead>
<tbody>
<tr>
<td>MUS MEAN</td>
<td>r=.40, p&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>MEM MEAN</td>
<td>r=.35, p&lt;.0002</td>
<td></td>
</tr>
<tr>
<td>N/Y/MUS N/RES</td>
<td>r=.38, p&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>START/MUS MEAN</td>
<td>r=.24, p&lt;.01</td>
<td></td>
</tr>
<tr>
<td>MUS N/RES</td>
<td>r=.36, p&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>MUS/SKILLS N/RES</td>
<td>r=.43, p&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>ATT N/RES</td>
<td>r=.50, p&lt;.0001</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Recapitulation of all correlations.

Legend: MUS – musicianship, MEM – memory, N/Y/MUS – number of years of musical training, START/MUS – the age of start of musical training, MUS/SKILLS – musical skills, ATT – attitude towards learning languages, MEAN – mean results, N/RES – number of correct responses.

Pearson correlation determines the extent to which values of two variables are related. Correlation coefficients can range from -1.00 to +1.00. The value of

...
-1.00 indicates a perfect negative correlation of two variables while a value of +1.00 indicates a perfect positive correlation. A value of 0.00 reveals a lack of correlation.

Examined variables are positively correlated, however, the correlations are small and moderate. Table 1 contains recapitulation of all correlations.

Comparison of mean results achieved by the two groups under research revealed that their performances differed significantly. Both Median and Mean values obtained by the two groups are close to each other, which means that there were not many residuals in both groups of examinees (Median 54, mean 53.74 for musicians and 37 and 38.17 for non-musicians, respectively).

In turn, more important differences between SD in non-musicians suggest that this group was not very coherent, and in the group were many subjects who were able to deal very well with the task but also many subjects whose performances were very poor. In addition IQR1 higher in non-musicians (39) suggested less coherence of the group and more differences between subjects’ performance.

<table>
<thead>
<tr>
<th></th>
<th>MUS</th>
<th>NONMUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>53</td>
<td>53</td>
</tr>
<tr>
<td>Mean</td>
<td>53.74</td>
<td>38.17</td>
</tr>
<tr>
<td>SD</td>
<td>16.71</td>
<td>23.11</td>
</tr>
<tr>
<td>SE</td>
<td>2.29</td>
<td>3.17</td>
</tr>
<tr>
<td>95% CI of Mean</td>
<td>49.130</td>
<td>58.341</td>
</tr>
<tr>
<td></td>
<td>to</td>
<td>31.801</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Median</th>
<th>IQR</th>
<th>95% CI of Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>MUS</td>
<td>54.00</td>
<td>26.00</td>
<td>to 63.00</td>
</tr>
<tr>
<td>NONMUS</td>
<td>37.00</td>
<td>39.00</td>
<td>to 51.00</td>
</tr>
</tbody>
</table>

Table 2. Comparative descriptive analysis of mean scores obtained by two groups of examinees.

3. CONCLUSION

Production scores obtained in the general analysis evidenced that musicians performed better than non-musicians in the whole experiment. The trained people were able to repeat more sentences and word sequences and with fewer errors.

An analysis of results achieved by subjects in the study in relation to subjects’ earlier musical expertise seems to confirm that music education affects second language acquisition and that the influence is not a myth but has thorough scientifically approved basis.

4. REFERENCES


---

1 The Inter-Quartile Range is a measure of the spread of or dispersion within a data set. The IQR is the width of an interval which contains the middle 50% of the sample, so it is smaller than the range and its value is less affected by outliers.