MUSICIANS OUTPERFORM NONMUSICIANS IN A STUDY WITH SHADOWING SPEECH

Barbara Pastuszek-Lipińska

Adam Mickiewicz University, Poznań, Poland
energin@wp.pl

ABSTRACT

In order to examine whether music education may be viewed as one of the factors which improve second language acquisition a research project was conducted.

Two groups of subjects : with musical training and without musical background were recruited and their responses to foreign language stimuli were recorded and then examined.

Research results provide an evidence that musically trained people encounter fewer difficulties in foreign language sounds discrimination.

Keywords: SLA, music, speech perception, speech production, memory.

1. INTRODUCTION

A number of studies revealed that a range of factors affects second language acquisition and several - not thoroughly examined - 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 cannot be omitted.

Only a limited number of studies examined a possible impact of music education on second language acquisition (e.g. [2, 4, 7, 8]). Moreover, there is still an ongoing discussion on the level of relationship. Namely, it is examined whether music education, music exposure or musicality improve human potential in language acquisition [5].

When looking for a possible transfer between musical training and language acquisition several approaches have been proposed. Most of them are based on the fact that training in music requires engagement and refinement of processes involved in the analysis of pitch patterns over time. Moreover the processes may be activated during interpretation of emotions conveyed by spoken utterances (cf. [2, 7, 8]). Indeed, some recent studies have provided evidence confirming the interrelation. Some of the processes are shared by both domains (e.g. processes that guide discrimination of emotional meaning conveyed by speech prosody and by music; acoustical cues shared by language and music), several of them are domain-specific. To date the issue has been discussed in several studies (e.g. [7, 8]).

For instance, Thompson, Schellenberg and Husain [7] in two of their experiments examined the hypothesis that music lessons generate positive transfer effects that influence speech perception. They provided evidence that musically trained participants outperformed untrained participants in extracting prosodic information from speech. On the basis of the finding the authors suggested the existence of a 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 study was developed to examine relationship of music education with foreign language acquisition. It was intended to cover some gaps that still exist in the field of inter-domain transfer examination and research on possible impact of music education on second language acquisition.

Amir and coworkers [1] have recently demonstrated that musicians had superior auditory skills. The authors proved that musicians even without vocal training showed greater vocal production accuracy in music.

On the ground of the finding it was intended to explore how musical training affects auditory perception and vocal production in language. The focus was given to sounds and construct perception and production.

The main goal of the study was to examine whether indeed active involvement in music influences second language acquisition.

2.1. The corpus

82 word sequences in 6 languages (English: American (15), British (14), Belgian Dutch (11),

French (12), Italian (10), Spanish: European (6) and from South America (4), and Japanese (10) were synthesized for the corpus. The ScanSoft® RealSpeak™ application was used for the purpose.

Languages were chosen according to the rules of typology. Among the word sequences are stimuli in both stress-timed, syllable-timed and moraes-timed languages. The corpus is composed of sentences (questions, statements, and orders), some phonological words, names and/or other short word sequences.

Thus, the stimuli differed phonemically, phonostylistically, and contained a variety of lexical items. The length of the sequences was also diversified. Table 1 contains examples from the languages used in the study.

Table 1: Exemplary word sequences.

Language	Example
American English	May I help you?
Belgian Dutch	Een fantastisch spektakel
British English	He looked nice and neat.
French	On va voir plus tard.
Italian	Molto rumore per nulla.
Japanese	Onegai-shimasu
European Spanish	Más vale tarde que nunca.
S.A.Spanish	Al tempo el consejo.

All stimuli were recorded on CD (3 times repeated with short gaps left between the repetitions of a word sequence, and a longer pause inserted so that to provide speakers with time needed to repeat a stimulus). Recorded corpus served for 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. They were not paid for the participation in the study. All subjects were aged from 15 to 69 years (mean 32).

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 exert negative impact on auditory perception).

2.4. Main procedure

In order to examine and compare how subjects with musical training and those without musical background perceive foreign language construct and sounds their ability to imitate speech stimuli was tested.

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 subjects. Examinees were not informed that they heard synthetic stimuli. All sentences produced by the subjects were recorded with Sharp MD-MT200 portable recorder and UNITRA-Tonsil Microphone MCU-53 with a linear characteristic.

The task was meant to examine the ability to integrate different components of linguistic information such as : phonology, syntax, and intonation. However, it was not a pure measure of the enumerated components. The main idea was to find a key to success or failure in language sounds and structure acquisition.

All recordings were heard and analysed by the author. The main goal was to determine whether all subjects performed at similar or different level.

2.5. Research results

Recordings were examined auditorily by the author in a randomized order. It was performed after a period of more than one year from the moment of data collection so that to ensure unbiased evaluation of all participants' production.

In the first round of data analysis the scoring procedure was based mainly on a general review and observation whether all speakers were able to repeat the speech material in the given time and with an appropriate accuracy.

In order to evaluate whether the task was not too difficult the Difficulty Factor, 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 optimal level of the factor equals 0.5.

The difficulty factor may be calculated using the following formula:

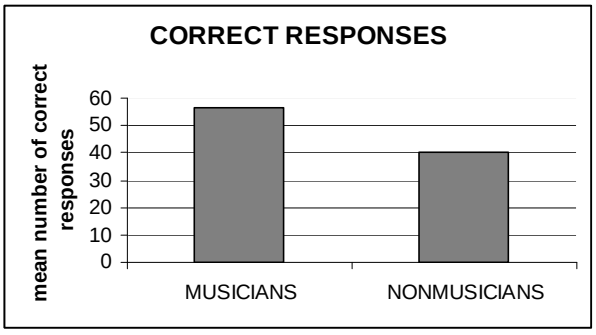
$$(1) \quad D = c / n$$

D - Difficulty factor
c - Number of correct answers
n - Number of respondents

As the purpose of the current study was to discriminate between different levels of performance, thus items with difficulty values between 0.3 and 0.7 would be most effective. The factor in the current study showed 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 nonmusicians, which means that the task was available for both groups of speakers.

Thus, it was not surprising that not all subjects were able to repeat all stimuli. The total percentage of correct repetitions (i.e. those very close to the original samples) was 45.95%.

Figure 1: Mean number of correct responses.



The group of musicians encountered fewer difficulties in shadowing speech repetition and produced 56.53% of correct responses to all the provided stimuli. As showed in Figure 1, the group of nonmusicians performed significantly worse than the group of musicians and produced 39.91% of correct repetitions. Moreover, most musicians performed better in all languages and all word sequences.

The obtained 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 minor problems with the task.

This observation corroborates with earlier findings that musically trained people, even without vocal training performed with greater vocal production accuracy [1].

It was noticed that number of correct productions differed among languages and among two groups of examinees. Namely, most musicians repeated all stimuli on time, however not all productions were fully faithful to the original.

2.6. More detailed data analysis

One of the aims of the study was to establish what types of errors were produced by participants. Therefore recorded productions were listened to and all inconsistencies, errors and mispronunciations produced by the speakers were analysed and assessed. Special attention was given to all mispronunciations that occurred systematically and in several speakers in the same words' or sounds' sequences.

Mispronunciations and inconsistencies were observed at both segmental and suprasegmental levels (cf. [6]). Errors of segmentation were mostly an important source of mistakes. It was determined that many subjects (more likely nonmusicians than musicians) changed several segments, repeated stimuli close 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 be an effect of phonemic restoration phenomenon observed by Kashino [3], among others. Namely, Kashino claimed that "the sounds we hear are not copies of physical sounds" and "what we perceive is the result of [an] unconscious interpretation" [3].

Mispronunciations observed in almost all the word sequences were generally: 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 neither whole words nor their parts (e.g. syllables, segments).

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 mentioned that different types of errors occurred in almost all participants' production. Speakers produced both errors of performance (slips of the tongue) and errors of competence (pronunciation) [6].

Moreover, in subjects' mimicry occurred many both native and foreign interferences.

2.7. Correlations and a 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 test of musical skills.

A moderate correlation was achieved between musicianship and mean results of subjects (r=.4, p<.0002). A partial correlation was obtained for results of memory test for music stimuli and mean results of subjects (r=.35, p<.0002). Similarly, a partial correlation was obtained between number of years of music education and performance expressed in the number of correct responses to the stimuli (r =.38, p<.0001). Furthermore, a partial correlation was obtained between the age of start of the music training and results achieved in the study (r=.24, p<.01). A partial correlation was also found between musicianship and the number of correct productions (r=.36, p<.0001). A moderate correlation was obtained between mean results of the test of musical skills and results achieved in the study (r=.43, p<.0001). The most significant correlation was obtained between the number of correct productions and attitude towards foreign language learning (r=.50, p<.0001). Interestingly, nonmusicians were slightly more susceptible to the variable (r=.50, p<.0001) compared with musicians (r=.48, p<.0002). Therefore, Pearson correlations are small and moderate.

Comparison of mean results obtained by musicians and nonmusicians revealed that their performances differed significantly. Both Median and Mean values obtained by the two groups were 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 nonmusicians respectively). In turn, the higher value of standard deviation (SD) in nonmusicians suggested that this group was not very coherent. Among the group were both subjects who managed very well with the task but also many subjects whose performances were very poor. In addition, the interquartile range (IQR), which was higher in nonmusicians, suggested incoherence of the group and more differences between subjects' performance.

Details of the statistical analysis are provided in Table 2.

Table 2. Comparative descriptive analysis of mean scores obtained by musicians (MUS) and nonmusicians (NONMUS).

	MUS		**NONMUS**	
N	53		53	
Mean	53,74		38,17	
SD	16,71		23,11	
SE	2,29		3,17	
95% CI of Mean	49,13	to 58,34	31,80	to 44,54

Median	**IQR**	**95% CI of Median**	
54,000	26,000	47,000	To 63,000
37,000	39,000	24,000	To 51,000

3. CONCLUSIONS

Production scores obtained in the general analysis evidenced that musicians performed significantly better than nonmusicians in the whole experiment.

Musically trained people were able to repeat more word sequences. Moreover, they produced fewer errors comparing to subjects without musical expertise. The general auditory analysis revealed that differences between musicians and nonmusicians persisted in all languages and in all word sequences.

The results of the study provide an evidence that music expertise indeed matters in second language acquisition.

4. REFERENCES

[1] Amir, O., Amir, N., Kishon-Riabin, L. 2003. The effect of superior auditory skills on vocal accuracy. *Journal of the Acoustical Society of America* 113 (2), 1102-8.

[2] Jacobson, L.S., Cuddy, L.L., Kilgour, A.R. 2003. Time tagging: a key to musicians' superior memory. *Music Perception* 20 (3), 307-313.

[3] Kashino, M. 2006. Phonemic restoration: The brain creates missing speech sounds. *Acoustical Science and Technology* 27 (6), 318-321.

[4] Magne, C., Schön, D., Besson, M. 2006. Musician children detect pitch violations in both music and lan guage better than nonmusician children: behavioural and electrophysiological approaches. *Journal of Cognitive Neuroscience* 18 (2), 199-211.

[5] Schellenberg, E.G. 2005. Music and cognitive abilities. *Current Directions in Psychological Science* 14, 322-325.

[6] Sobkowiak, W. 2001. *English Phonetics for Poles.* Poznań: Wydawnictwo Poznańskie.

[7] Thompson, W.F., Schellenberg, E.G., Husain, G. 2003. Perceiving prosody in speech effects of music lessons. In: Avanzini, G., Faienza, C., Minciacchi, D., Lopez, L., Mayno, M. (eds), *The Neurosciences of Music.* New York: The New York Academy of Sciences, vol. 999, 530-532.

[8] Thompson, W.F., Schellenberg, E.G., Husain, G. 2004. Decoding speech prosody. Do music lessons help? *Emotions* 1, 46-64.