Ewa Kusz

Musical Intelligence and its impact on English pronunciation skills in the process of second language acquisition.

The strong need to define the concept of "intelligence" started in antiquity, whereas the first ideas of objective measures of human intellect emerged at the end of the 19th century, when Alfred Binet attempted to establish the types of cognitive abilities that could separate the normal children from the abnormal, and to measure such differences. As a matter of fact, there are various perspectives with regard to the nature of intelligence and the ways it should be addressed. It is possible to distinguish at least two types of approaches: first, there are researchers who believe that there is such a thing as "general intelligence" or, in other words, one unified and general factor that may be defined as the core of human intellect, as it influences intelligent performance. This theory was firstly proposed by the psychologist Charles Spearman (1863-1945) who observed that there was a positive correlation between various intelligence tests that attempted to investigate cognitive skills. Thus, he suggested that it is possible to experimentally measure intelligence as a general cognitive ability.

On the other hand, there are scholars who perceive intelligence as a wide variety of skills and attitudes that make individuals differ from one another in their ability to comprehend complex ideas, to learn new things or to adapt effectively to the environment. On the basis of the above, this group of researchers believe that it is possible to observe completely different and independent types of intelligences. One of the first scholars who proposed that intelligence should not be considered as a general ability but rather as a set of abilities was Thurstone. It was also Gardner, the father of the theory of Multiple Intelligences (MI), who strongly criticised the ideas presented by psychometricians, including Spearman's theory of general intelligence.

Since linguistic and logical-mathematical intelligences were always more in the spotlight and the centre of researchers' attention, this study focuses mainly on the concept of musical intelligence and ways of recognising individuals with musical talent who have a good sense of chord analysis, pitch change or tonal and rhythmic memory. It is not without reason that Gardner's theory of Multiple Intelligences was successfully adopted in the field of education, as it suggests that teachers should on the one hand pay greater attention to their students' natural predispositions, and on the other devote more time to students' weaknesses in order to improve them. Although the MI theory encountered a lot of criticism, mostly alleging that is was based on an intuition rather than experimental procedures and that the supposed “intelligences” were nothing more than personality traits, it has gained great popularity over the last decades. The major reason for that is probably the occurrence of postulates which state that each type of intelligence might constitute separate sets of skills that can develop independently, and as a result they might be dissociated from general intelligence. It seems rational to assume that although linguistic intelligence is naturally relevant to second/foreign language acquisition, there is also musical intelligence that should be taken into consideration. It is the link between music and language that continues to intrigue researchers in terms of making learning easier and faster. Thus, in line with the perspective proposed by Gardner, the process of learning a second language can be facilitated if one makes optimal use of their musical intelligence.

When one is concerned with the measurement of musical intelligence, a series of fundamental questions arise: *Is musical intelligence a unitary ability or multi-dimensional one? If yes, how many sub-components constitute it? Is musical intelligence normally distributed in the population or is it rather ‘all-or-none’ ability? Are there groups of people who are more musically intelligent than others? Is it possible to distinguish factors that enable us to measure musical intelligence?* It is with these questions in mind (and many others that will be revealed later) that the present study has been projected and written. This will help to make the reader aware of the complexity of the concept of musical intelligence and its measurement. Since some of the questions formulated above clearly reach beyond language, it should not come as a surprise that this work, while adopting the interdependence between musical intelligence and L2 pronunciation proficiency as its primary object of research, will attempt to incorporate a number of phenomena which are usually associated with other disciplines, such as psychology, psycholinguistics, statistics and even philosophy. In terms of second language acquisition (SLA) research, there is one central claim which states that there is a specific talent for learning foreign languages, which enables us to distinguish between talented and untalented learners. However, language is a part of one’s identity which belongs to a person’s whole social being. Thus, the process of second language learning involves far more than learning skills. It also involves the ability to adopt new social and cultural behaviours, as well as ways of being.

Following that, it is not easy to carry out studies in the research field of musical intelligence and its impact on SLA, especially in terms of L2 pronunciation ability. This is partly because it is difficult to state whether this specific aspect of second language acquisition is influenced by such cognitive psychological constructs as intelligence or whether it is merely language specific in nature. Additionally, the complexity of selecting the appropriate measurement approach, as well as measurement instruments, has been very challenging, as was the fact that various combinations of the selected measures often provide mixed results, which may cause problems in the correct interpretation of the findings. The following study has the objective of providing a comprehensive examination of the convergence between musical intelligence and pronunciation ability of a second language from various perspectives. On the one hand, it considers psychological influences of musical intelligence on L2 pronunciation talent, but most importantly it has the objective of finding general insights into the nature of both musical intelligence and L2 pronunciation skills, and it intends to explore more specific interactions between the examined parameters that may have an impact on the above-mentioned relationship.

Bearing the above in mind, the present work has been divided into five chapters. The first, *The concept of intelligence and approaches to its study*, intends to provide an organising and integrating framework for the chapters in the remaining parts. Thus, it raises epistemological issues concerning various concepts of intelligence and it provides a rather comprehensive history of the general field of intelligence and attempts at measuring it. Also, it focuses on delineating the nature of human intellect as it relates to and is defined by society and culture. The second chapter moves into a detailed description of the theory of Multiple Intelligences proposed by Howard Gardner with particular reference to musical intelligence which, according to the founder of the theory, "allows people to create, communicate, and understand meanings made out of sound". This chapter also constitutes an attempt at describing the relationship between musical intelligence and other intelligences in order to relate this type of human intellect to broader contexts. Additionally, this part of the study is largely concerned with the issue of measuring musical intelligence, thus the most popular musical intelligence tests (Seashore's, Wing's and Gordon's) are reviewed. The last two subsections of the second chapter delve into the relationship between musical intelligence and the process of second language acquisition, especially in terms of pronunciation skills. It must be mentioned that there are some unforeseen challenges in describing L2 pronunciation assessment due to the piecemeal contributions of individual researchers on the one hand, and a strong need to develop toward acceptance of the inevitability of the use of automated speech recognition technology on the other. Chapter three presents a literature review regarding the convergence between musical intelligence and second language pronunciation skills, as well as the assessment of speaking. A detailed description of the prosodic features of speaker accentedness is the "prelude" to the methodology of this research, which provides information on the number of participants, the types of instruments utilised in the study and finally the experimental procedure. Among a number of areas of phonetic research, the fourth chapter, *Research results*, takes into consideration some of the most problematic issues, including the convergence between musical intelligence and pronunciation skills, and secondly, the assessment of pronunciation. The first part of this chapter focuses mainly on the results obtained from Wing's musical intelligence test, whereas the second pays attention to the outcomes of the pronunciation test (measured both by three independent native speakers and by *Praat*). Last but not least, chapter five delves into a detailed discussion and a comprehensive interpretation of the obtained results and empirical contributions which will confirm or refute the relationship between the level of musical intelligence and L2 pronunciation ability.

Wing's musical intelligence test results

The analysis of the results obtained from Wing's musical intelligence test has shown that in all three tasks the data distribution is non-symmetric, yet most of the results centre around the average score (there is a minimum skewness to the left in tasks 1 and 2, whereas in task 3 the data is skewed to the right). For tasks 1 and 2, the average scores are higher than the median (averages of 8.48, 17.18, and medians of 8 and 16 for tasks 1 and 2 respectively) which suggest that the students of English Philology tend to gain worse results more frequently than the overall average score for chord analysis and detection of pitch change. Conversely to tasks 1 and 2, the results obtained in task 3 revealed that the majority of respondents gained better scores than the average result (hence, the data is skewed to the right and the median (18) was higher than the average (17.95)). It should not come as a surprise that the overall results confirm non-symmetric data distribution which is minimally skewed to the left (average 43.61 and median 42). The norms established for the Polish students presented in either centile or sten scale enabled us to compare their results to the results obtained by a group of people between 20-25 years old. The findings show that there is a tendency for the sample of students of English Philology to gain a higher level of musical intelligence than the group at the same age who do not study in this faculty.

Pronunciation test results – native speakers' assessment of students' utterances

The results of this study have shown that the relationship between Wing's musical intelligence test results and native speakers' assessment in pronunciation tests is considered statistically significant. In terms of the results that were obtained only in the pronunciation test that was assessed by the panel of native speakers, it should be noted that although there are numerous examples in the research literature where rating scales have been strongly criticised, a panel of English native speakers recruited as expert raters for this research succeeded in obtaining perceptual judgments of English Philology students' pronunciation for comprehensibility, intelligibility and fluency. The raters listened to the speech samples in a computer-administered rating task recorded with the use of *Praat* software. Also, all the speech samples were normalised for peak and mean amplitude. After listening, English native speakers rated the answers provided by students using ten-point scales. To check for rating consistency, several statistical tests were computed. Given high inter-rater reliability, it was possible to acquire final rating scores and use them as measures of overall L2 pronunciation ability.

To continue the establishment of a complete picture of the data set, the table below shows the overall mathematical averages of the population divided with regards to Wing's musical intelligence test with level distinction (low, medium and high), as well as the pronunciation level evaluated by the three English native speakers. Taking into consideration the average score value, which equals 5.19 points, it was decided to divide the assessment given by the natives into two groups, namely above average and below average.

	Above average	**Below average**	**Total**
Low	7	20	27
	7.45%	21.28%	28.72%
Medium	10	26	36
	10.64%	27.66%	38.30%
High	18	13	31
	19.15%	13.83%	32.98%
Total	35	59	94
	37.23%	62.77%	100.00%

The data set presented above shows the number of English Philology students divided into low-, medium- and high-level groups regarding the results of the musical intelligence test. They were presented in juxtaposition with the above average and below average levels which represent the results of the pronunciation test that the participants in the research obtained. There is a clear dependence of the scores gained in the pronunciation test and the students' level of musical intelligence. Although the majority of the population (62.77%) obtained scores that were lower than the average (5.19 points), it is immediately obvious when we examine the frequency chart that the higher one's level of musical intelligence, the higher the level of L2 pronunciation proficiency is likely to be.

The results of automatic measurement of L2 pronunciation proficiency with the use of *Praat*

The major aim of this part of the research is an attempt to answer the question of whether the level of musical intelligence measured in the sample group of students correlates with the results achieved in the pronunciation test obtained from *Praat*. As was indicated in sections 4.1, 4.2, and 4.3, musical intelligence is identified with the results achieved by the students in Wing's test (the number of points and stens). The musical intelligence test enabled us to distinguish three groups of participants in terms of the scores they gained, namely low, medium, and high.

A description of the principles and definitions applied for the collection of statistical data on the students' pronunciation test results obtained from *Praat* is the first step to the detailed statistical analysis of this part of the research. Hence, we shall start with a review of such standard statistical measures as average, median, standardised kurtosis, as well as standardised skewness. The last two variables enable us to provide information regarding whether the data are normally distributed or not. This is extremely important especially in terms of the use of the parametric test ANOVA, for which normally distributed data is the first condition that must be met in order to use it. The normality of the distributions will also be tested by Shapiro-Wilk and Kolmogorov-Smirnov tests.

As previously mentioned, there are five variables obtained from the *Praat* script for the automatic measurement of students' pronunciation skills, namely:

- F_0 Range;
- Speech Rate;
- Articulation Rate;

- Average Syllable Duration (ASD);
- Pause Duration;

A description of the above-mentioned variables, as well as a literature review on the automatic measurement of pronunciation skills which justifies the use of such temporal measures in terms of pronunciation, were set out in chapter 3. Additionally, it should be noted that for each of the variables the *Praat* script generated a model value which was based on a paradigm presented in the *Spelling-to-sound* exercise in *Longman Pronunciation Dictionary*. Following that, the results that 94 students of English Philology achieved in the pronunciation test were compared to the model value and then the standard deviation was calculated. The smaller the standard deviation was, the better pronunciation test result the students achieved. The statistical description and analysis of this part of the study focus mainly on the absolute value which is the margin between the above-mentioned variables and their paradigms (model values). The results obtained from the statistical analysis were grouped into five categories and named as follows:

- F_0 Range Difference (F_0 D);
- Speech Rate Difference (SR D);
- Articulation Rate Difference (AR D);
- Average Syllable Duration Difference (ASD D);
- Pause Duration Difference (PD D);

The correlation between the level of musical intelligence obtained from Wing's musical intelligence test and pronunciation test (results of which were obtained from *Praat*) should first be verified by certain statistical tests. They are divided into two categories: non-parametric (sometimes called distribution-free tests) and parametric tests. The Kruskal-Wallis test is one of the non-parametric tests that was used in this research to analyse the results provided by native speakers in the pronunciation test. Unlike parametric tests, non-parametric statistics are based on fewer assumptions. For instance, it is not required for a data set to be normally distributed. This assumption is required for the use of parametric tests, for which the sample data must come from a population that follows a probability distribution based on a fixed set of parameters. ANOVA (Analysis of variance) is an example of such a parametric test that was used in this study in order to determine whether there are any statistically significant differences between the means of independent (unrelated) groups (low, medium, and high). However, before ANOVA is applied in this research, it is of utmost importance to provide four conditions that the data set must meet to use the analysis of variance method:

1. Normality – the distributions of the residuals are normal (each group sample is drawn from a normally distributed population);
2. Homogeneity of variances – the variance of data in the studied groups must be very similar or the same;
3. Factor effects are additive (the data must be numeric);
4. Independence of observations – the results obtained from the studied groups cannot influence each other, and the observations are sampled randomly, independent of each other;

Regarding the results of research which were obtained from *Praat*, it is immediately obvious that assumptions 3 and 4 met the conditions for ANOVA. All the results are numeric, and sampled independently of each other. It is also worth noting that the results gained by the participants in one group (e.g. 'low') did not influence other groups. Additionally, the students of English Philology did not communicate among one another while they were attending Wing's musical intelligence test. Similarly, the pronunciation test was also taken individually, as the students were asked to do it at home in a stressfree environment and e-mail their responses.

As far as assumptions 1 and 2 are concerned, their applicability in the research results requires deeper analysis based on appropriate statistical tests. Following that, the Shapiro-Wilk and Kolmogorov-Smirnov tests will be used in order to test normality for the sample population, whereas the homogeneity of variances will be tested with the use of Bartlett's test.

As stated above, one of the main criteria that decides which type of test (parametric or non-parametric) should be applied in the process of analysing the impact of one factor on the results gained by the sample population (English Philology students) is to test the normality of its distribution. Thus, in order to determine the validity of statistical measures, each of the five results of the pronunciation test obtained from *Praat* is conducted in terms of the normality of its distribution.

	Wing Level			**Total**
Statistic	**Low**	**Medium**	**High**	
Count	27	36	31	94
Average	120.848	93.9244	70.6968	93.9977
Median	120	92.25	61.4	91.2
Standard Deviation	55.5047	53.6125	47.4481	55.3131
Minimum	8	0.2	2.9	0.2
Maximum	291.9	230.1	213.1	291.9
Range	283.9	229.9	210.2	291.7

Lower Quartile	85.4	52.45	35.2	50.7
Upper Quartile	140.5	113.45	99.6	125.1
Standardised Skewness	2.1255	1.5738	2.4828	3.2149
Standardised Kurtosis	3.1148	0.54810	1.5862	2.2515

Table 1. F_0 Range Difference - statistical results.

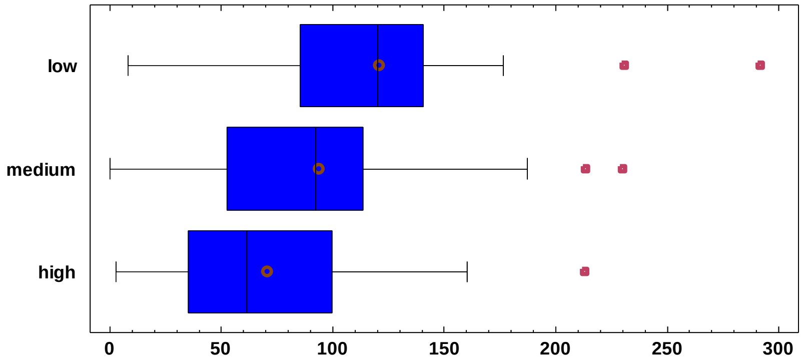
The results presented in Table 13 refer to the first variable that was tested with the use of *Praat*, namely F_0 Range Difference. As indicated above, they refer to the level of deviation that differs from the model value. In other words, the lower the values are, the better results were achieved. Comparing both arithmetic means and medians in terms of musical intelligence level (low, medium, high), it can easily be noticed that the values achieved by the students with the highest level of musical intelligence are the smallest (average 70.69, median 61.4). Consequently, the values for the 'medium' group of respondents are higher (average 93.92, median 92.25). What is interesting here is that the values for the last group (low) are notably the largest (average 120.84, median 120). Similarly, the values for standard deviation also show the smallest differences between the model value and the results of the 'high' group (47.44) slightly larger for the 'medium' group (53.61), and the largest for the 'low' group (55.5).

This may suggest that there is a significant dependence between the results obtained from Wing's musical intelligence test and F_0 Range Difference. Nonetheless, in order to verify the validity of this assumption, the normal distribution of the sample population should be tested, as the analysis of standardised skewness and standardised kurtosis allow us to assume that there might be some deviations from the normal distribution, especially among the 'high' and 'low' groups, as well as the total results. To verify this, the Shapiro-Wilk test of normality was used.

Test	***Statistic***	***P-Value***
Shapiro-Wilk W	0.9559	0.0129

Test of normality for F_0 Range Difference (94 results/respondents).

As presented above, since the smallest p-value amongst the tests performed is greater than or equal to 0.05, we cannot reject the idea that F_0 Range Difference comes from a normal distribution with 95% confidence. Nonetheless, it is worth checking whether there are no outliers in a random sample from a population. The outliers can be easily identified in a box-and-whisker plot.



F_0 Range Difference by Wing level.

The above information is of great interest as it demonstrates that there are outliers in each group: low, medium and high. The outliers are marked as red, squared dots outside the whiskers of the boxes. It means that they are markedly different from other values in a sample population. Interestingly, regardless of the group, the values of the outliers are significantly higher (not lower) than the ‘typical’ values achieved by the students of English Philology. Following that, the next step is to verify whether the outliers have an impact on the distribution of the sample group. Thus, the next step is to remove these values from the study. Moreover, this action will be repeated in terms of the latter values obtained from *Praat* which will be discussed in the next subsections.

	Wing Level			**Total**
Statistic	**low**	**medium**	**High**	
Count	25	34	30	89
Average	0.3425	0.349	0.2567	0.3161
Median	0.289	0.2325	0.241	0.244
Standard Deviation	0.2688	0.3604	0.1994	0.2887
Minimum	≈0*	0.032	0.006	≈0*
Maximum	0.902	1.725	0.885	1.725
Range	0.902	1.693	0.879	1.725
Lower Quartile	0.115	0.154	0.088	0.128
Upper Quartile	0.513	0.422	0.316	0.443
Standardised Skewness	1.1446	5.7779	2.8524	8.1781
Standardised Kurtosis	-0.5034	7.8448	2.369	12.78

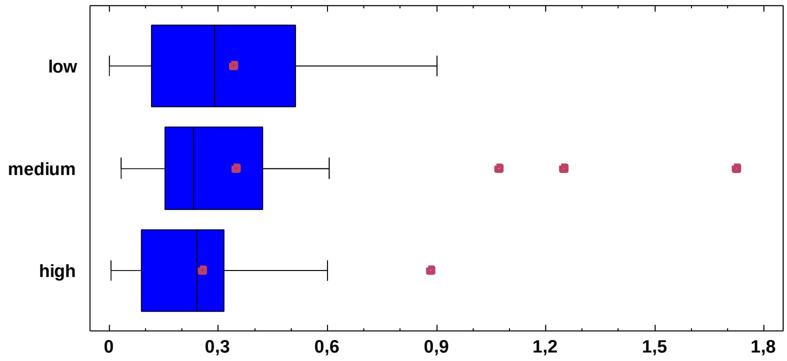
Speech Rate Difference – statistical results (89 results).

Next, we shall turn to the values obtained from the pronunciation test (measured by *Praat*) in which 89 students of English Philology participated. The results shown in Table 15 refer to Speech rate Difference (SR D) and they indicate that it is not possible to state whether there is a significant correlation between the level of musical intelligence and the Speech rate results obtained from *Praat*. As expected, if we compare the average scores, we can easily notice that the smallest value is in the 'high' group. However, it is surprising that the average score for the 'medium' group is higher than in the 'low' one. This suggests that the students whose level of musical intelligence is the lowest, scored better results in the pronunciation test than the students in the 'medium' group, even though the latter's musical intelligence is at a higher level. The comparison of medians provides even more interesting outcomes. The highest median value is for the 'low' group, which unsurprisingly demonstrates that the greater the margin between the results obtained from *Praat* and the model value, the less impact musical intelligence has on pronunciation skills. However, the median value for the 'medium' group of students presented in Table 15 is smaller than the median for the 'high' group, which suggests that English Philology students with a medium level of musical intelligence gained better results in measuring Speech rate. Similarly to the previous variable, the Speech rate difference was tested in terms of its normal distribution.

Test	***Statistic***	***P-Value***
Shapiro-Wilk W	0.8265	≈0*

Test of normality for Speech Rate Difference (89 results).

The values of standardised kurtosis and standardised skewness suggest that the results are not normally distributed. This is confirmed by the results of the Shapiro-Wilk test which shows that the p-value is outside the range of -2 to +2 and less than the significance level of 0.05. This indicates significant departures from normality and tends to invalidate any statistical test regarding standard deviation. Following that, the null hypothesis of the normal distribution of the sample population should therefore be rejected.



Speech Rate Difference from Wing level (89 results).

Similarly to the results of F_0 Range Difference, the analysis of the box-and-whisker-plot which presents the results of Speech rate Difference (SR D), shows that there are some outliers which should be excluded from the data in order to ascertain the validity of the results. There are four students whose results are outside the range of the box-and-whisker plot (the values are significantly higher than the results that the rest of the students achieved). As a result, the further study will focus on the remaining 85 students of English Philology. Before we move to the analysis of the results of the next variable, it is worth focussing on each box presented in Figure 11. There is a significant difference between the quartiles for the low, medium, and high groups. As the medians (the line that divides each box into two parts) and the averages (red dots) have already been discussed, we shall now take a closer look at the upper and lower quartiles. If we reject the outliers in the 'medium' and 'high' groups, it is worth noting that the widest range of results appears in the 'low' group, whereas the 'high' group has the smallest range of values. Also, the lower quartile in the 'high' group equals only 0.088, whereas for the 'medium' and 'low' groups the lower quartiles represent significantly higher values: 0.15 and 0.11 respectively. We shall now focus on the next variable – Articulation Rate Difference.

The standard deviation for all three groups presented in the table below shows that the best results in Articulation Rate Difference were achieved in the 'high' group (students with the highest level of musical intelligence), in which the difference between the model value of standard deviation and the one that students gained was the smallest. Cognately, the average score for the 'high' group has the lowest value, which means that the difference between the model average score and the one that the most musically intelligent respondents achieved is the smallest. The ‘medium’ group gained a slightly higher average score, whereas the ‘low’ one showed the largest discrepancy between the average values. However, we may come to a different conclusion when we look at the medians. It was the ‘medium’ group that achieved the smallest difference between its result and the model value. The participants with the highest level of musical intelligence (the ‘high’ group) gained worse results than the ‘medium’ group. As was expected, the largest discrepancy between the model value and the results of Articulation Rate Difference (almost twice as high as in the ‘medium’ group) was found in the ‘low’ group.

	Wing Level			
Statistic	**low**	**medium**	**high**	**Total**
Count	25	31	29	85
Average	0.4163	0.3479	0.3057	0.3536
Median	0.398	0.193	0.252	0.252
Standard Deviation	0.3318	0.3218	0.2211	0.2944
Minimum	0.009	0.027	≈0	≈0
Maximum	1.209	1.096	0.881	1.209
Range	1.2	1.069	0.881	1.209
Lower Quartile	0.093	0.092	0.179	0.106
Upper Quartile	0.62	0.512	0.483	0.564
Standardised Skewness	0.9541	2.4026	1.3863	3.2226
Standardised Kurtosis	-0.5112	0.0734	0.1172	0.0982

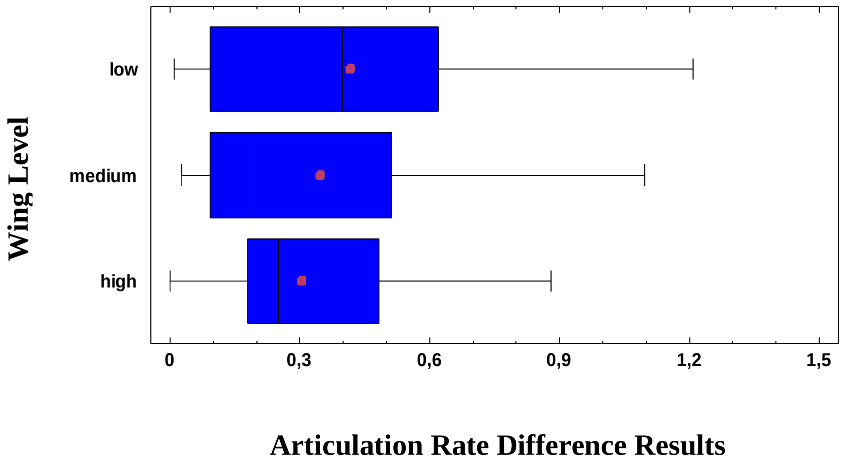
Articulation Rate Difference - statistical results (85 results).

Based on the analysis of these two variables (average score and median), the following may be stated for the Articulation Rate Difference results. As was expected, the largest deviations from the model value appeared in the group of students with the lowest level of musical intelligence. This may indicate that the higher the score achieved by the English Philology students in Wing’s musical intelligence test, the better the scores gained in the pronunciation test measured by *Praat*. However, it is not possible to draw that conclusion unless a test for a normal distribution is applied. Hence, we shall now use the Shapiro-Wilk test to analyse the normal distribution of the sample to ascertain the validity of the results.

Test	***Statistic***	***P-Value***
Shapiro-Wilk W	0.9013	≈0

Test of normality for Articulation Rate Difference (85 results).

The data set provided in Table 18 indicates that we can reject the idea that Articulation Rate Difference is normally distributed with 95% confidence, as the smallest p-value amongst the tests performed is less than 0.05 ($p<0.05$). The medians and quartiles of the Articulation Rate Difference results are visualised in the box-and-whisker plot (Figure 12).



Articulation Rate Difference by Wing level.

The analysis of the previous box-and-whisker plots which presented the data set for F_0 Range Difference and Speech Rate Difference, as well as the lack of normal distribution in Articulation Rate Difference results, indicated that there might be some outliers for Articulation Rate Difference values. Surprisingly, as presented in Figure 12, there are no outliers for this data set. More importantly, the results have shown that average scores are not significantly differentiated from each other. The values are situated within the upper and lower quartiles (they do not appear outside the boxes) and they are also based on the results gained from Wing's musical intelligence test. Also, the range of values is the smallest for the 'high' group and the largest for the 'low' group, as was expected.

The next variable that was obtained from *Praat* in the automatic measurement of pronunciation ability is Average Syllable Duration (ASD). Table 19 presents the statistical deviation values (ASD Difference) between the results that the students of English Philology achieved and the model variables. It is worth noting that the average scores for the students whose musical intelligence level was classified as high and medium, are very similar (to be more precise, the better result was gained by the 'medium' group). The greatest discrepancy between the average score and the model value can be noticed in the 'low' group, which unsurprisingly is in line with expectations.

Statistic	Wing Level low	medium	high	Total
Count	25	31	29	85
Average	0.0288	0.02032	0.0217	0.0233
Median	0.03	0.01	0.02	0.02
Standard Deviation	0.0247	0.0202	0.0154	0.0203
Minimum	0	0	0	0
Maximum	0.12	0.08	0.05	0.12
Range	0.12	0.08	0.05	0.12
Lower Quartile	0.01	0.01	0.01	0.01
Upper Quartile	0.04	0.03	0.03	0.03
Standardised Skewness	4.0879	3.1144	0.5725	6.1067
Standardised Kurtosis	7.1716	1.7002	-0.7339	9.7208

ASD Difference - statistical results (85 results).

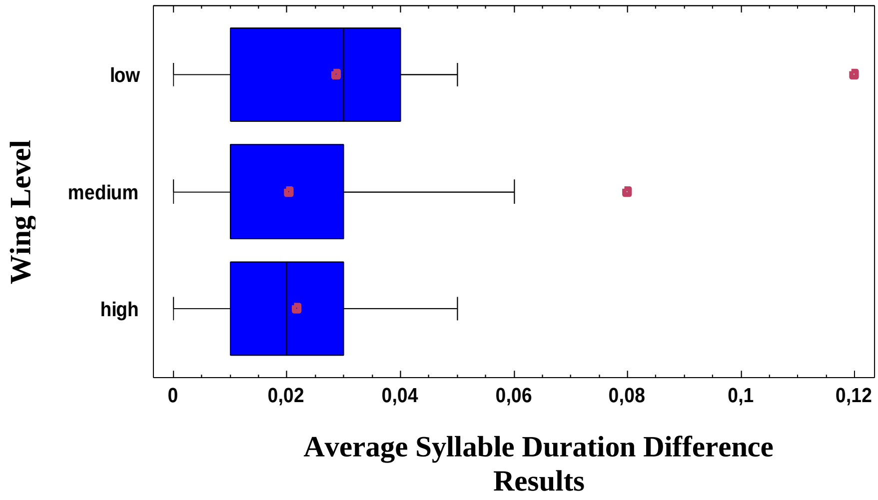
While comparing median values for ASD Difference, it should be noted that the highest score (which in this case shows the greatest deviation from the model value) was gained by the students whose level of musical intelligence was the lowest. A more surprising result, however, is the fact that the median value for the 'medium' group is even two times lower than for the 'high' group. This suggests a weak correlation between the results obtained from Wing's musical intelligence test and ASD deviations from the model value.

Cognately to the analysis of the previous variables, we shall now apply the Shapiro-Wilk test of normality in order to ascertain the validity of the ASD Difference data set.

Test	***Statistic***	***P-Value***
Shapiro-Wilk W	0.8687	≈0

Test of normality for ASD Difference (85 results).

On the basis of the above, we can conclude that the data set for the Average Syllable Duration Difference obtained from *Praat* is not normally distributed (Table 20), as the p-value amongst the tests performed is less than 0.05 ($p<0.05$). Following that, we shall now focus on the analysis of the box-and-whisker-plots which describe low, medium and high groups in terms of their Average Syllable Duration Difference results.



ASD Difference by Wing level.

The analysis of the box-and-whisker plots for ASD Difference results enables us to distinguish and then exclude outliers for ‘medium’ and ‘low’ groups. As a result, there are 83 remaining results gained from the students of English Philology that will be interpreted in terms of the next variable obtained from *Praat*. However, before we do that, we need to examine the boxes presented in Figure 13. Interestingly, the lower quartile of the ‘low’ group covers the range of both quartiles for the ‘medium’ and ‘high’ groups. Additionally, if we exclude the outlier in the ‘medium’ group, the range of the values for the ‘medium’ and ‘high’ groups will be the same. The final variable that is taken into consideration is Pause Duration Difference.

Comparing the Pause Duration Difference results gained by the remaining 83 students to the Pause Duration model values revealed that it is the ‘low’ group that achieved the greatest values. Similarly to the previous parts of the study, this means that the least musically intelligent participants scored the worst results in the pronunciation test measured by *Praat*. The average values for the ‘low’ and ‘high’ groups were very close to each other. As expected, the median value gained by the most musically intelligent students was the lowest, whereas it was slightly higher for the ‘medium’ group and the highest for the ‘low’ group. Yet, Table 21 shows that the averages for the ‘high’ and ‘medium’ groups are not as they were expected, specifically the average for the ‘high’ group is slightly larger than the one for the ‘medium’ group. More importantly, the results of standard deviation, range, standardised kurtosis and skewness strongly suggest that the data set for Pause Duration Difference is far from normal (i.e. it is not normally distributed).

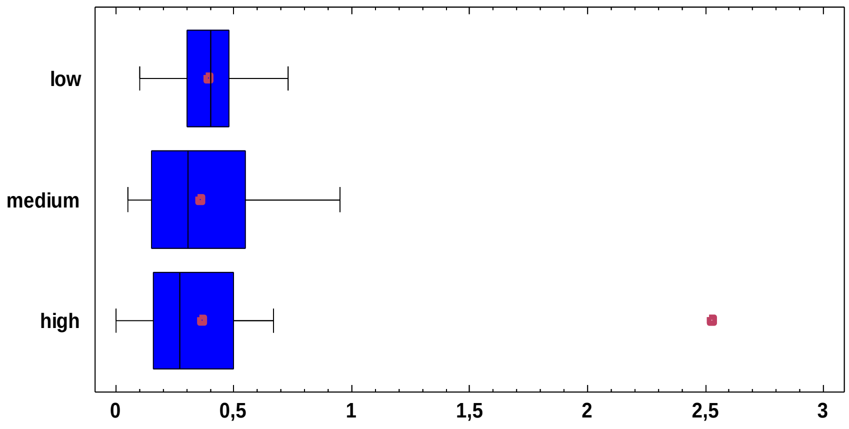
	Wing Level			Total
Statistic	**Low**	**Medium**	**High**	
Count	24	30	29	83
Average	0.3942	0.3607	0.3666	0.3724
Median	0.4	0.305	0.27	0.32
Standard Deviation	0.16	0.2471	0.4583	0.3173
Minimum	0.1	0.05	0	0
Maximum	0.73	0.95	2.53	2.53
Range	0.63	0.9	2.53	2.53
Lower Quartile	0.3	0.15	0.16	0.17
Upper Quartile	0.48	0.55	0.5	0.5
Standardised Skewness	0.5776	1.686	8.7006	14.7432
Standardised Kurtosis	-0.084	-0.26881	20.591	47.3406

Pause Duration Difference - statistical results (83 results).

Furthermore, it is worth noting that the Pause Duration Difference values for the students whose musical intelligence level was the highest ranged more than the values of F_0, Speech rate, Articulation rate or Average Syllable Duration within the same group. This can easily be noticed in terms of the maximum and minimum values. Contrary to expectations, the greatest deviation from the model Pause Duration Difference value appeared in the 'high' group. The difference between the 'low' and 'medium' groups was insignificant: 0.73 seconds for low, 0.95 seconds for medium and 2.5 seconds for the 'high' group. The ongoing analysis of the Shapiro-Wilk test and box-and-whisker plots will determine and verify the validation of the above-mentioned issue.

Test	***Statistic***	***P-Value***
Shapiro-Wilk W	0.7275	0

Test of normality for Pause Duration Difference (83 results).



Pause Duration Difference by Wing level.

The results obtained from the Shapiro-Wilk test indicate that the null hypothesis should be rejected, as the discussed data set is not normally distributed. One of the main reasons for the lack of normality, as well as significant deviations noticed in the 'high' group, could be the outliers. To verify the validity of this assumption, the box-and-whisker plots for all three groups (low, medium, and high) were produced. There is one outlier in the 'high' group which should be eliminated from the data set. On the basis of this, there are 82 remaining results in terms of the pronunciation test (24 in the 'low' group, 30 in the 'medium' group, and 28 in the 'high' group) which will be taken into consideration for further analysis.

It is extremely important to stress the fact that the elimination of outliers from the data set was a necessary step in view of investigating the interdependence between the results of the pronunciation test measured by *Praat* and Wing's musical intelligence test scores. The removal of outliers enabled us to discover which results might have affected the overall correlation between the final scores of the musical intelligence test and the results gained in the automatic pronunciation test obtained from *Praat.* Also, the occurrence of outliers in the data set turned our attention to the recordings provided by students in terms of the pronunciation test. Thus, the recording of each student whose result achieved in the automatic pronunciation test turned out to be an outlier was listened to again. In most cases, the recorded answers were disrupted by the students' repetitions, falters or exaggerated pauses.

The above analysis also demonstrates that there is a significant dependence between the results obtained from Wing's musical intelligence test and native speakers' assessment pronunciation test results, as well as the *F_0* range that was obtained from *Praat*. The lack of correlation between the level of musical intelligence and Speech Rate, Articulation Rate, Average Syllable Duration, as well as Pause Duration may suggest that either some different variables should be taken into consideration for future research or there is still a need to investigate the nature of this relationship. When considering the multiple linear regression models presented in this chapter, it clearly emerges that they can be used for predicting and analysing the interdependence between the above-mentioned variables.