ANALYSIS OF LANGUAGE ATTITUDES BY SOCIO-DEMOGRPHIC INFORMATION - II

CONSULTANT: SHUYI TAN

As a continuation of socio-demographic analysis - I, we will focus on the following variables in this analysis:

- the gender of respondents
- the race of respondents
- the country/countries in which the respondent was growing up
- the number of languages respondents speaks
- the number of years studying French

How the pattern of language attitudes differ by these factors will be investigated.

1. Effect of Speaker Accents

This section will examine how differently respondents classified by various socio-demographic variables evaluated five types of French accents respectively. To minimize the effect of speaker race, the scores given to recordings recorded by speakers of the same accent will be summed up for each respondent.

1.1. Gender. First of all, we are interested in to what extent participants of different gender evaluated various French accents differently. It can be observed from Figure A.1 and A.2 that the evaluation scores of most accents given by two genders are close except that on average the L2 French accent seems to be less favored by male respondents than female respondents. To check if the gender differences are significant, two-sample Wilcoxon rank-sum tests are implemented. The Wilcoxon rank-sum test is a non-parametric alternative to the two-sample t test which is based solely on the order in which the observations. It is more flexible than the two-sample t test because it does not assume the data have a known distribution. The null hypothesis of the Wilcoxon test is that two populations have the same distribution with the same median. If we reject the null, it indicates we have evidence that one distribution is shifted to the left or right of the other. With the results presented in Table 1, it reveals that there is no significant gender differences existing in any of the five French accents.

Speaker Accent	Test Statistic	P-value
Acadian	2127.5	0.345
African	2258.5	0.560
European	2207.0	0.474
L2 French	2012.5	0.190
Québec	1978.5	0.153

TABLE 1. Results of Wilcoxon Tests on Gender for Each Accent

1.2. **Race of Respondents.** Besides gender, the effect of respondent races on language attitudes is also of interest. In the questionnaire, respondents were asked to voluntarily identify their ethnicity before listening to recordings. Exploring whether listeners' ethnicity were associated with their ratings will help to provide insights into the hierarchy of language attitudes.

1.2.1. ***Grouping by if the respondent are white or not***. Based on the answers of self-identified ethnicity, there are 34 white respondents and 59 non-white respondents. We can observe from Figure A.3 and Figure A.4 that except for the Québec accent, white participants tend to give higher scores to most accents compared to non-white participants. However, it is revealed by the results of Wilcoxon tests in Table 2 that none of the differences between two ethnicity groups in each speaker accent is significant. In other words, whether being ethnically white did not result in significantly different languages attitudes.

Speaker Accent	Test Statistic	P-value
Acadian	822.0	0.150
African	833.5	0.177
European	907.0	0.446
L2 French	939.5	0.615
Québec	1121.5	0.346

TABLE 2. Results of Wilxocon Tests on Respondent Ethnicity (white vs. non-white) for Each Accent

1.2.2. ***Grouping by Three Major Race Groups***. Considering the lack of statistical power due to the diversity presented in the races of respondents, respondents who have revealed their races but are neither white nor Chinese will be classified as "others." In general, respondents are classified into three major race groups: white (34), Chinese (22), and others (37). It is displayed in Figure A.5 and Figure A.6 that the evaluation patterns are alike in the groups of Acadian accent, African accent and European accent. In these three accents, the language attitudes of Chinese respondents and white respondents appeared to be slightly more positive compared to the group of "others." When it comes to the accents of L2 French and Québec, the ratings of three race groups are very close. The results of Kruskal Wallis tests summarized in Table 3 verifies our empirical findings: no significant difference is found in any group of speaker accent.

Speaker Accent	Test Statistic	P-value
Acadian	2.09	0.352
African	1.93	0.381
European	1.53	0.465
L2 French	0.26	0.879
Québec	1.00	0.607

TABLE 3. Results of Kruskal-Wallis Tests on Major Race groups for Each Speaker Accent

1.3. **Grown-up Country/Countries.** The language exposure from the places growing up may potentially shape people's language attitudes. Therefore, this subsection will check the countries in which respondents were growing up. To have a more thorough examination on the effect, the countries growing up will be grouped in multiple ways. Out of 153 respondents who have answered this question, 108 of them grew up in Canada. Classifying the respondents by the number of grown-up

countries, 140 grew up in one country, among whom 99 grew up in Canada, 25 grew up in Asia, 10 grew up in America, 5 grew up in Europe, and 1 grew up in Africa. For those whose who grew up in more than one countries, 9 of them grew in both Canada and some other country/countries.

1.3.1. ***Grouping by if Growing up in Canada.*** Overall, 70% of respondents were growing up in Canada and 30% were not. To see if growing up in Canada influenced respondents' language attitudes toward different speaker accents, the mean values of scores for each speaker accent are presented in Figure A.8. Along with the medians shown in A.7, we can observe that respondents growing up in Canada tend to have more positive attitudes towards most accents than those who were not except for the European accent. To confirm whether the differences between the group of growing up in Canada and the group of not growing up in Canada are significant, two-sample Wilcoxon tests were implemented again. Observing the p-values in Table 4, except that the p-value of L2 French accent is relatively small (close to 5%), no other statistical significance was found to exist, which aligns with the empirical observations. In this case, we can say that respondents growing up in Canada held nearly significantly more positive attitudes to L2 French than those who were not.

Speaker Accent	Test Statistics	P-value
Acadian	2027.5	0.1070
African	2179.5	0.3170
European	2484.5	0.8290
L2 French	1944.5	0.0521
Québec	2412.5	0.9460

TABLE 4. Results of Wilxocon Tests on Whether Growing up in Canada for Each Accent

1.3.2. ***Grouping by if Growing up in More than One Countries.*** Considering that growing up in more than one country may have cultural impacts on people's language attitudes, we also classify the participants based on whether they were growing up in more than one country. It is demonstrated in Figure A.9 and Figure A.8 that participants who lived in more than one country when growing up apparently gave higher scores to all accents. We can leave aside the extremely close mean values shown in the European accent in Figure A.10 because the mean is sensitive to outliers. As evidenced by Table 5, participants with the background of multi growing-up countries rated the accents of Acadian and L2 French significantly more positively. The effect of growing up in multiple countries are less significant in the accent of African and Québec, and it is totally not significant when the speaker accent is European.

Speaker Accent	Test Statistics	P-value
Acadian	503.0	0.00779
African	623.5	0.06120
European	843.0	0.66300
L2 French	443.0	0.00226
Québec	619.0	0.05720

TABLE 5. Results of Wilxocon Tests on Whether Growing up in More Than One Country for Each Accent

1.3.3. ***Grouping by Country/Continent.*** In addition, for those who were not growing up in Canada, they are also grouped by continents: 25 grew up in Asia, 10 grew up in the Americas, 5 grew up in Europe, and 1 grew up in Africa. Continent Americas is defined to be a landmass comprising the totality of North and South America except for Canada in this analysis. Since there is only one respondent growing up in Africa, he/she was removed from the visualizations and tests. Figure A.11 depicts how participants growing up in different continents/countries evaluated a certain French accent and Figure A.12 compares the mean ratings of respondents across five accents. We can tell from two graphs that participants from the Americas (excluding Canada) gave the highest or second-highest scores to the most accents compared to participants growing up in other places. However, the L2 French accent is least favored by people who grew up in the Americas (excluding Canada). For the European accent and the Québec accent that we are particularly interested in, it is revealed by the plots that the European accent is most favorable to respondents who grew up in Americas and Europe, while the Québec accent is most appreciated by the Americas (including Canada). The outcomes of the Kruskal Wallis Tests are summarized in Table 6. Based on the p-values, we can see that statistical significance only shows up in the group of European accent. Since the output of the Kruskal-Wallis test only tell that there is a significant difference between groups, but which pairs of groups are different remains unknown. Therefore, Dunn's test is implemented for post-hoc analysis. However, after adjusting the p-values for multiple comparisons, no statistical significance was found in the pair-wise comparisons of country/continents in the group of European accent. The contradictory results between Wilcoxon test and post-hoc analysis may occur because we do not have enough statistical power as a result of groups having small sizes. When pairwise comparison tests are not statistically powerful, it is less likely to detect significant differences.

Speaker Accent	Test Statistics	P-value
Acadian	7.57	0.1090
African	4.04	0.4010
European	9.59	0.0479
L2 French	5.86	0.2090
Québec	4.01	0.4030

TABLE 6. Results of Kruskal-Wallis Tests on Grew up Country/Continent for Each Speaker Accent

1.4. **Number of Language(s) Spoken.** For the number of languages respondents speak, the majority of them speak three languages (82; 53%), which is followed by two languages (45; 29%), four languages (24; 15%), and more than five languages (5; 3%). To explore the relation between language attitudes and the number of languages a respondent can speak, the distribution of evaluation scores is visualized using jitter plots along with smoothed regression lines. A jitter plot is a variant of the strip plot with a better view of overlapping data points, used to visualize the distribution of many individual one-dimensional values. It is demonstrated in Figure A.13 that the evaluation scores are slightly negatively associated with the number of languages spoken when the speaker accent is Acadian or African. Otherwise, the regression lines in another three accents are almost horizontal. To quantify the relations, correlation analysis was conducted for each speaker accent. Since the variable *number of language(s) spoken* is ordinal, we will use Spearman's rho to assess the association, which is a non-parametric alternative to Pearson's correlation. Spearman's correlation coefficient ranges from -1 to 1. The larger the absolute value of the coefficient is, the stronger the relation between the variables would be. An absolute value of 1 indicates a perfect relation, and a value of zero indicates the absence of an ordinal relation. For the p-values coming with the correlation, it checks whether the result is meaningful or is just down to chance. The results of the correlation analysis are summarized in Table 7. As evidenced by Table 7, no statistical

significance was found to exist in all groups of speak accents. We may not draw conclusions from these correlation coefficients since they are hardly meaningful.

Speaker Accent	Spearman's Rho	P-value
Acadian	-0.086	0.289
African	-0.099	0.220
European	-0.035	0.666
L2 French	-0.014	0.865
Québec	0.123	0.127

TABLE 7. Results of Spearman Correlation between Language Attitudes and Number of Language(s) Spoken

1.5. **Years of Studying French.** For the years of studying French, it ranges from 2 to 25 among all survey participants. From Figure A.14, we can see that the ratings are positively associated with the year of learning years in the groups of Acadian accent, African accent, and European accent, and the upward trend is most apparent when the speaker spoke in European accent. Otherwise, the relation in the groups of L2 French accent and Québec accent is not evident. To quantify the effect of year of studying French on language attitudes, the Pearson correlation coefficient will be calculated, which measures a linear dependence between two variables. Having checked the normality of evaluation scores and the years of studying French, the results of Pearson correlation for each speaker accent are shown in Table 8. There was a statistically significant but weak correlation in the European accent with a small p-value, from which we can infer that people who have learned French longer tend to hold more favorable attitudes towards the European accent.

Speaker Accent	Pearson's Rho	P-value
Acadian	0.049	0.544
African	0.068	0.402
European	0.192	0.016
L2 French	-0.009	0.909
Québec	-0.036	0.657

TABLE 8. Results of Pearson Correlation between Language Attitudes and the Years of Studying French

Furthermore, to amply the effect of studying years, we may group the data on a five-year scale. Excluding one respondent who did not answer, 37.8% of respondents have learned $5 - 10$ years (59), 26.3% have learned for $0 - 5$ years (41), 24.4% for $10 - 15$ years (38), and 11.5% for over 15 years (18). Before perform correlation analysis, four groups were encoded into rankings as follows: $0 - 5$ years as 1, $5 - 10$ years as 2, $10 - 15$ years as 3, and over 15 years 4. The results of Spearman correlation coefficients and p-values are summarized in Table 9. Interestingly, no statistically significant correlation was found between the evaluation accents and the years of studying French on a five-year scale. Only the p-value of the European accent is close to $\alpha = 0.05$. Performing a subsidiary correlation analysis for the European accent using the Kendall correlation, it returns a correlation coefficient of 0.044 with a p-value 0.056. In this case, we can say that there exists a weak correlation but mostly statistically significant between the years of learning French on five-year intervals and language attitudes towards the European accent. The more five years people learn French, the more positive was their language attitudes towards the European accent.

Speaker Accent	Spearman's Rho	P-value
Acadian	0.071	0.380
African	0.104	0.195
European	0.155	0.053
L2 French	0.056	0.486
Québec	-0.035	0.664

TABLE 9. Results of Spearman Correlation between Language Attitudes and the Years of Studying French on Five-year Interval

2. **Effect of Speaker Race**

The section will investigate if there exists potential relation between the speaker race and language attitudes of the respondents. Following the same logic of the analysis of speaker accents, the ratings of recordings recorded by speakers of the same race will be summed up for each respondent.

2.1. **Examine Languages Attitude towards A Certain Race Grouped by Categorical Socio-demographic Variables.** For categorical socio-demographic variables such as gender and countries where respondents grew up, we will initially explore the distribution of the evaluations scores with visualizations and examine whether various groups rates recordings by speakers of a certain race significantly differently using statistical tests (Table 10) . Following the same order of analysis as sections 1.1 and 1.2, we start with the gender of respondents. It is shown in Figure A.15 that the median of scores given by male respondents are higher than those of male respondents when the speaker is black. Conversely, recordings by white speakers received more positive feedback from females respondents. When it comes to the countries they grew up, if we group the respondents by whether they were growing up in Canada, Figure A.16 shows that the patterns of two race groups are alike. Similarly, the grouped boxes of two races in Figure A.17 are alike when classifying the participants by whether they were growing up in more than one country. As for the grouping of country/continent they growing up, it can be observed from Figure A.18 that there is no obvious difference between two race groups.

To consolidate these empirical findings, Wilcoxon tests are recruited for the analysis of gender and Kruskal-Wallis tests are performed on variables related to countries in which they grew up. It is demonstrated in Table 10 that statistically significant difference is only found in the group of black speakers when grouping the respondents by whether they were growing up in more than one country, which reveals that respondents with the background of growing up in multiple countries held significantly more positive attitudes towards black speakers compared to those who were not. However, the difference was not significant when the speaker is white. Overall, the outcomes reveal that only the factor of growing up in multiple countries significantly influenced respondents' language attitudes towards black speakers.

Speaker Race	Grouping Variable	Test	Test Statistics	P-value
Black	Gender	Wilcoxon	2227.5	0.987
	Growing up in Canada	Wilcoxon	2089.5	0.173
	Growing up in Multiple Countries	Wilcoxon	393.5	0.000733
	Country/Continent Growing up	Kruskal-Wallis	7.251	0.123
	Ethnicity of Respondent (White vs. Non-White)	Wilcoxon	980.0	0.858
	Ethnicity of Respondent (White vs. Chinese vs. Others)	Kruskal-Wallis	0.08	0.961
White	Gender	Wilcoxon	2033.0	0.428
	Growing up in Canada	Wilcoxon	2132.0	0.233
	Growing up in Multiple Countries	Wilcoxon	682.0	0.137
	Country/Continent Growing up	Kruskal-Wallis	4.345	0.361
	Ethnicity of Respondent (White vs. Non-White)	Wilcoxon	827.5	0.163
	Ethnicity of Respondent (White vs. Chinese vs. Others)	Kruskal-Wallis	2.86	0.240

TABLE 10. Results of Statistical Tests on Categorical Grouping variables across Race

2.2. Correlation Analysis of Socio-demographic Numeric Variables on A Certain Race.

For numeric variables, correlations with two-tailed tests of significance were performed to investigate the relation with speaker race and language attitudes. Following the same methodology as sections 1.3 and 1.4, the correlation coefficients and p-value are presented in Table 11, from which we can see that there is no significant relation existing between language attitudes and either the number of languages spoken or the year of learning French.

Variable of Interest	Race	Correlation Type	Rho	P-value
Number of Languages spoken	Black	Spearman	-0.013	0.875
	White	Spearman	-0.033	0.680
Year of Learning French (Continous)	Black	Pearson	0.079	0.332
	White	Pearson	0.055	0.498
Year of Learning French (On five-year interval)	Black	Spearman	0.103	0.203
	White	Spearman	0.074	0.360

TABLE 11. Result of Correlation Analysis on Numeric Socio-demographic Variables across Race

2.3. Comparison of Language Attitudes Over Race.

After checking whether respondents of different groups within a grouping variable evaluated speakers of a specific race differently, in this subsection, we will examine if respondents from a certain group held different attitudes towards speakers of dichotomous races. For each group, the mean value of scores given to speakers of different races are calculated, and in the meantime, we conduct Wilcoxon tests to validate whether participant rates speakers of different races significantly differently. We have learned from previous analysis that overall recordings of black speakers are more favored by the respondents compared to those of white speakers. Comparing the average evaluation scores of two speaker groups (μ_{Black} and μ_{White}) in Table 12, we can see that the preference for black speakers exists in most groups

except for the group of respondents who grew up in Europe. On average respondents growing up in Europe presents more favorable attitudes towards white speakers than black speakers, which is not statistically significant nevertheless. Now We check other pair-wise comparisons in Table 12 variable by variable. Regarding gender, both groups gave significantly higher scores to black speakers. For the ethnicity of respondents, non-white respondents rates black speakers significantly higher, while the preference is not significant in the group of white respondents. Further grouping the respondents into "white", "Chinese", and "others", significantly more favorable attitudes towards black speakers only appear in the group of "others." When it comes to where respondents grew up, it is shown that respondents who grew up in Canada held significantly more positive attitudes towards black speakers, while speakers of different races did not receive significantly different evaluations from respondents who grew up outside of Canada.

	μ_{Black}	μ_{White}	Test Statistic	P-value
Gender				
Man	208.54	195.13	986.0	0.024
Woman	207.43	198.72	7633.5	0.023
Race of Respondent				
White	206.29	204.85	630	0.528
Non-White	209.17	197.92	2141	0.031
Race Group of Respondent				
White	206.29	204.85	630.0	0.528
Chinese	210.05	201.68	282.0	0.354
Others	208.65	195.68	875.5	0.039
Grew up in Canada				
True	209.88	199.94	7065.0	0.007
False	202.07	195.00	1151.5	0.264
Grew up in Multiple Countries				
True	234.77	210.46	128.0	0.027
False	205.06	197.37	11403.5	0.018
Country/Continent Growing up				
Canada	207.68	198.34	5905.5	0.013
America	205.60	200.50	56.5	0.650
Asia	197.44	189.52	366.5	0.299
Europe	180.00	203.20	9.0	0.548

TABLE 12. Results of Wilcoxon Signed Rank Tests on Speaker Race

Appendix A. Figures
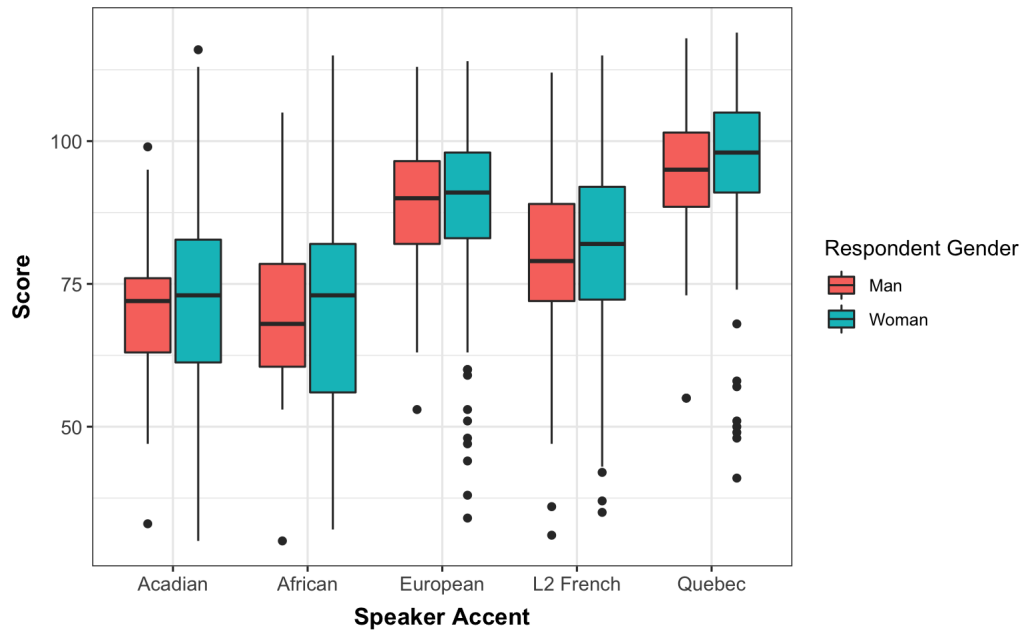
FIGURE A.1. Scores by respondent gender and speaker accent
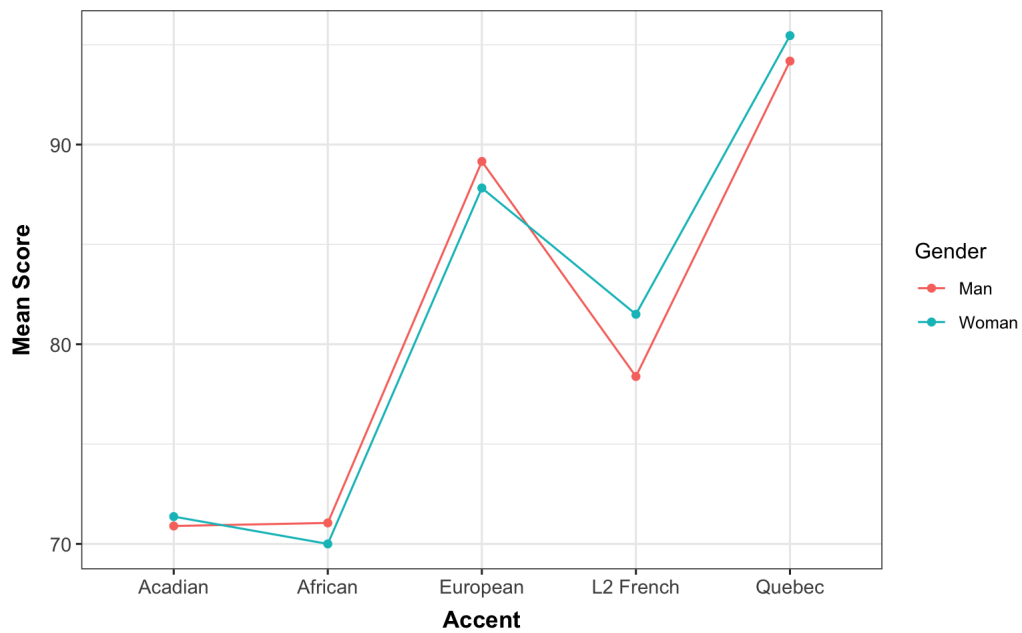
FIGURE A.2. Mean score of respondents by gender across each speaker accent
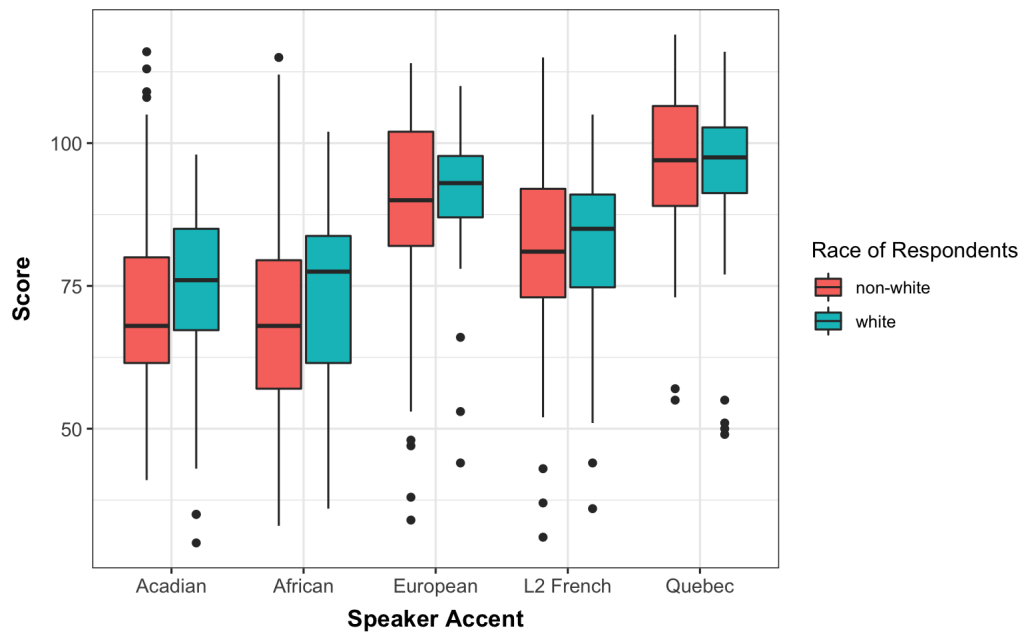
FIGURE A.3. Scores by respondent ethnicity and speaker accent
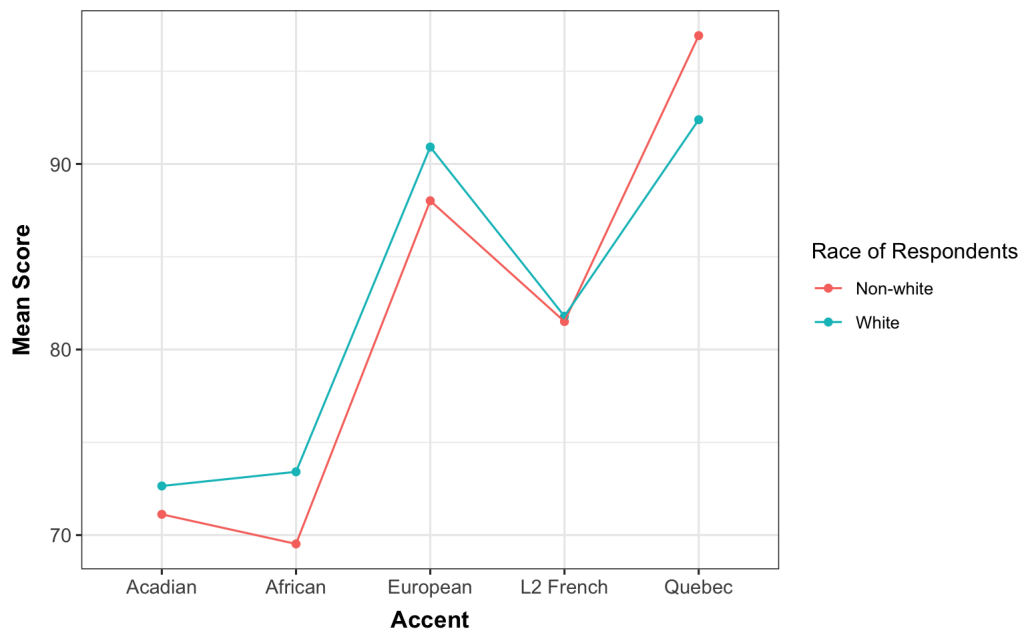
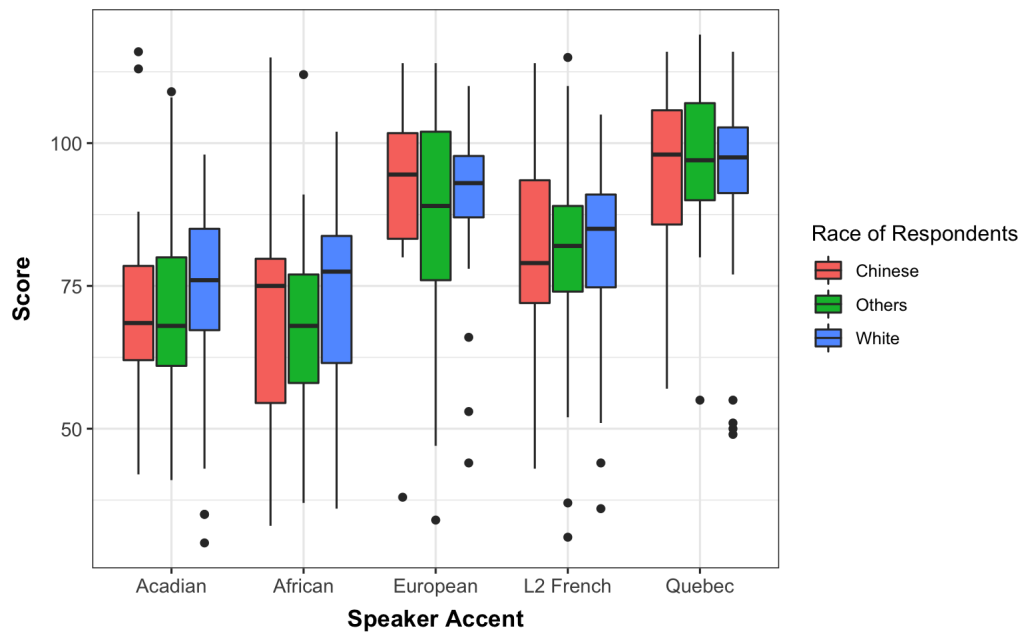FIGURE A.4. Mean score by respondent ethnicity across each speaker accent

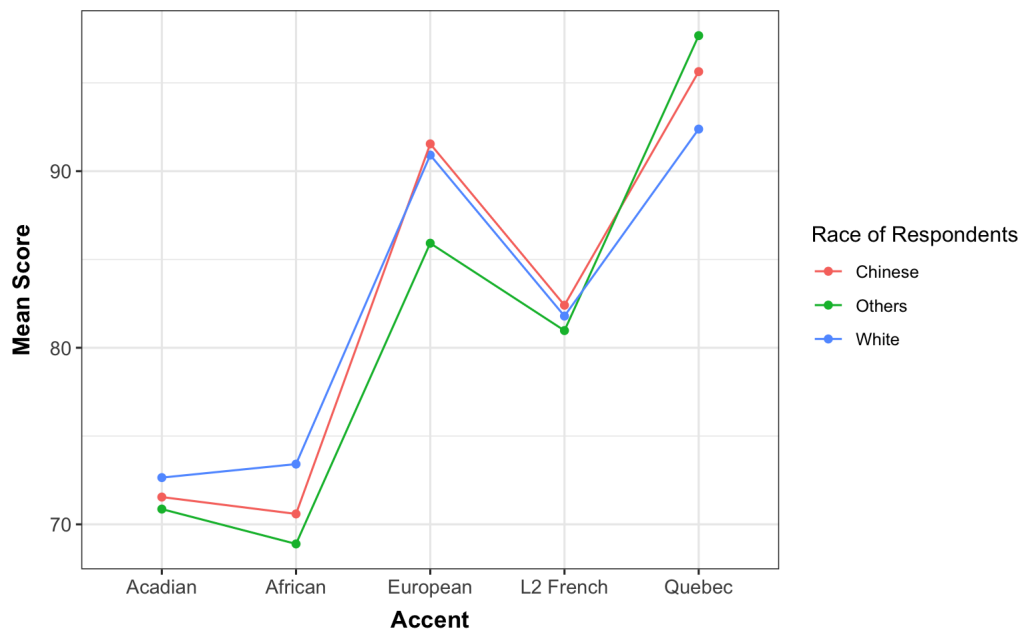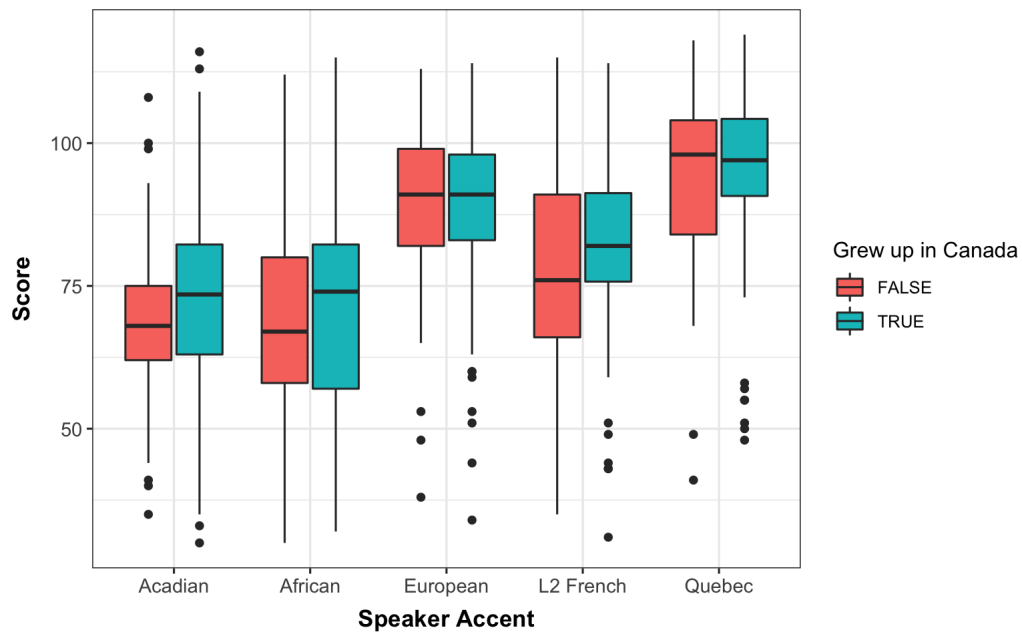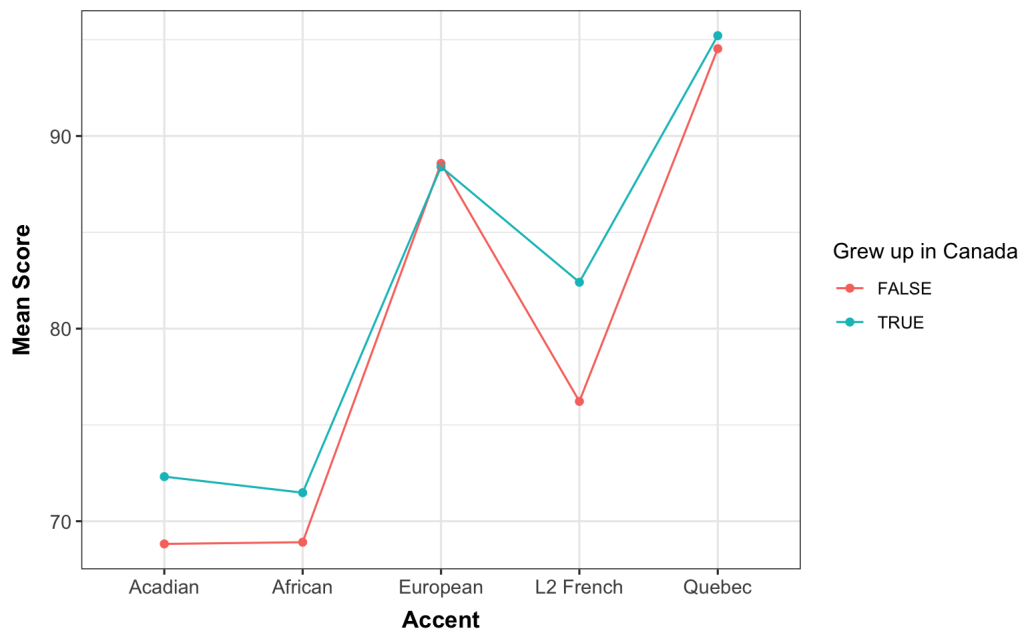FIGURE A.5. Scores by respondent ethnicity group and speaker accent

FIGURE A.6. Mean score by respondent ethnicity group across each speaker accent

FIGURE A.7. Scores by grown-up country and speaker accent

FIGURE A.8. Mean score of respondents by grown-up country across each speaker accent
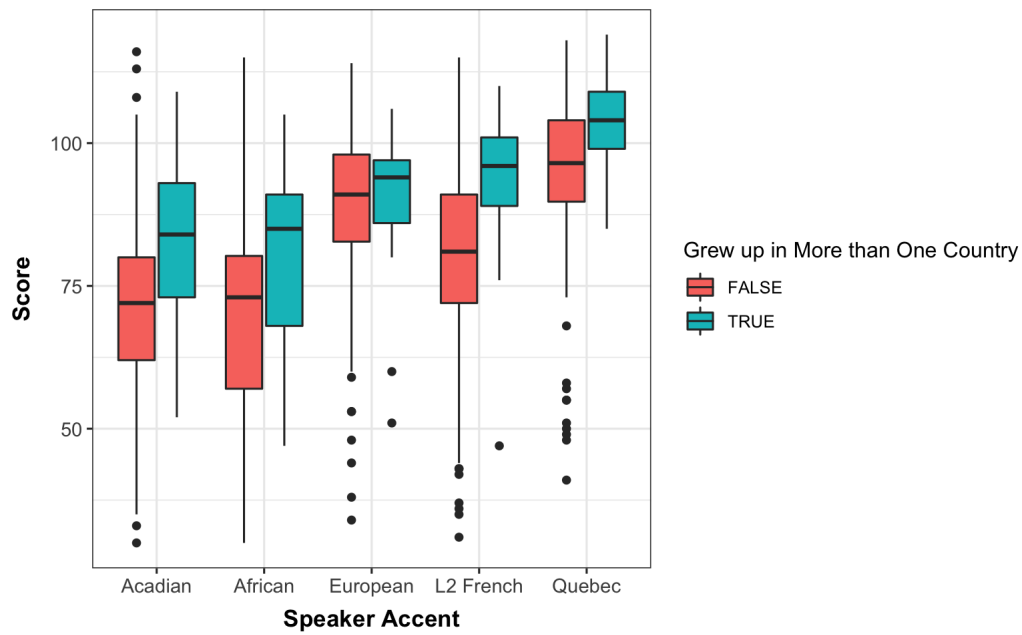
FIGURE A.9. Scores by the number of grown-up countries and speaker accent
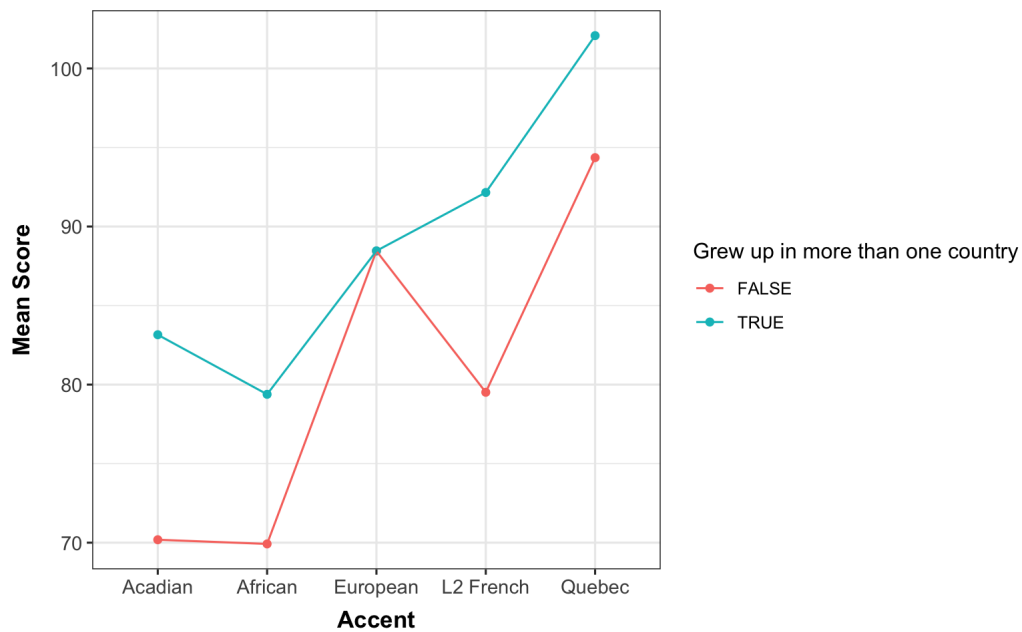
FIGURE A.10. Mean score of respondents by the number of grown-up countries across each speaker accent
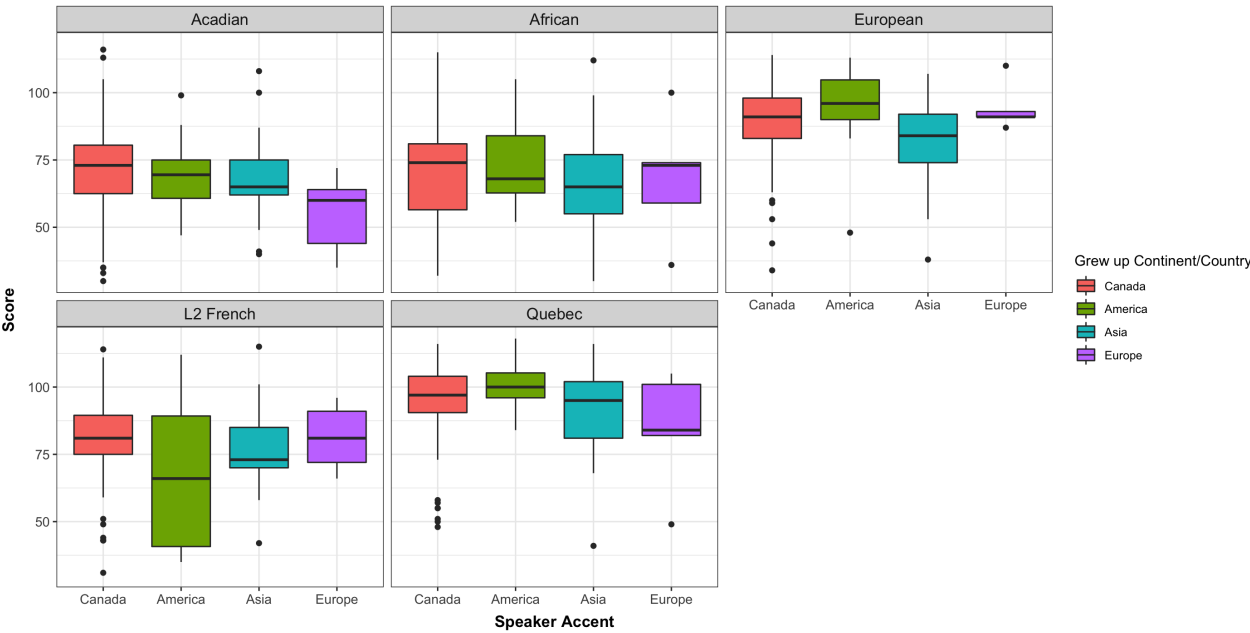
FIGURE A.11. Scores by grown-up country/continent and speaker accent
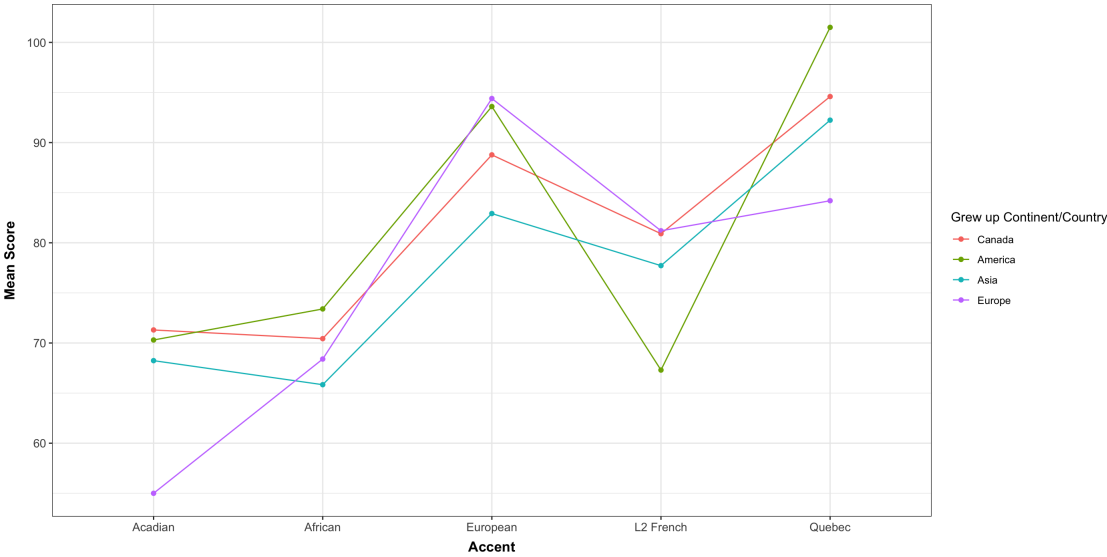
FIGURE A.12. Mean score of respondents by country/continent across each speaker accent
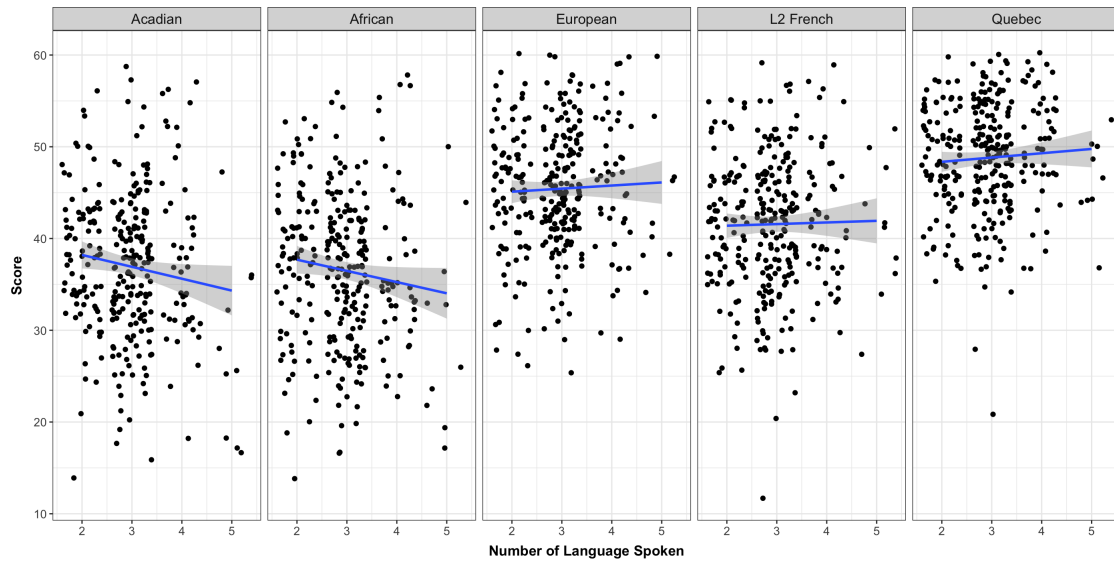
FIGURE A.13. Evaluation score of respondents vs. the Number of Languages Spoken across each speaker accent
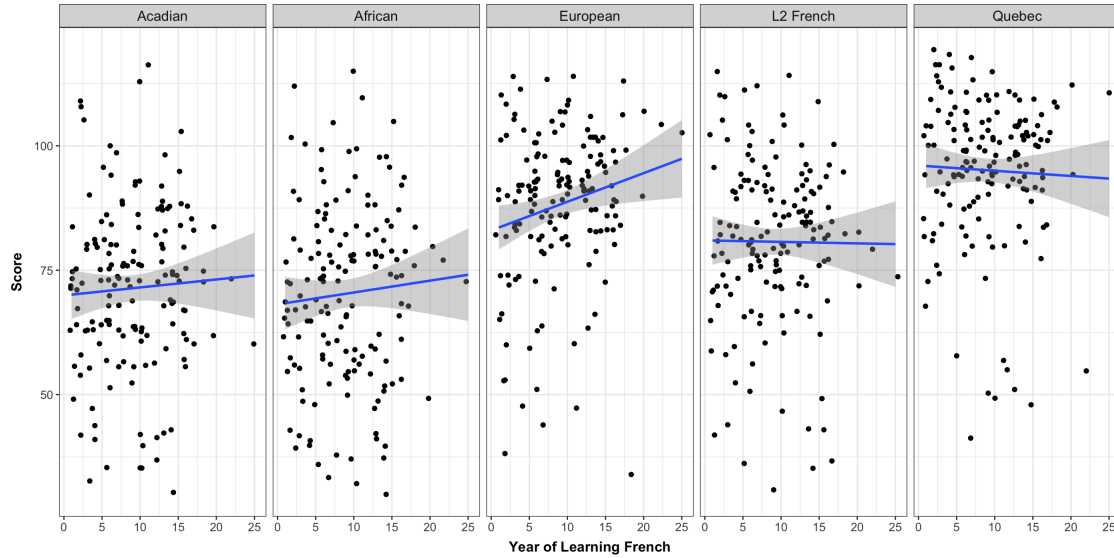
FIGURE A.14. Evaluation score of respondents vs. Years of Studying French across each speaker accent
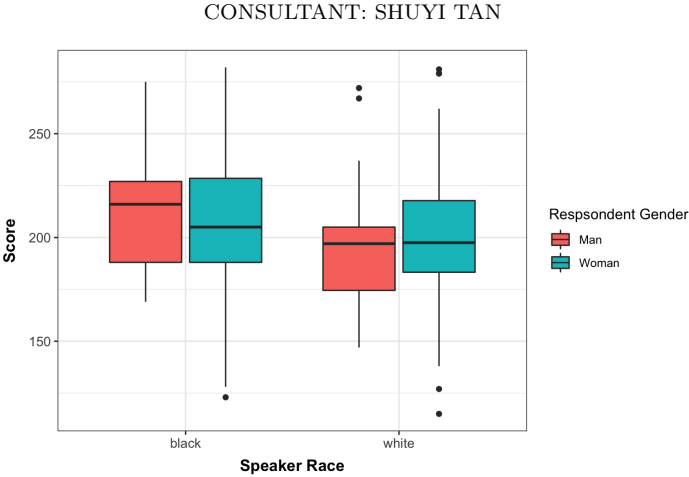
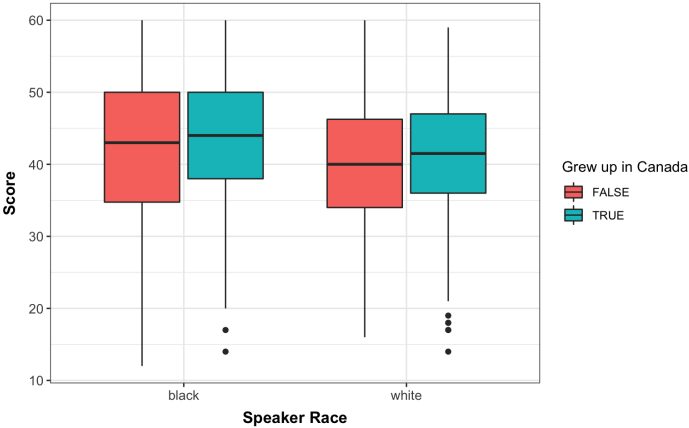FIGURE A.15. Scores by respondent gender and speaker race

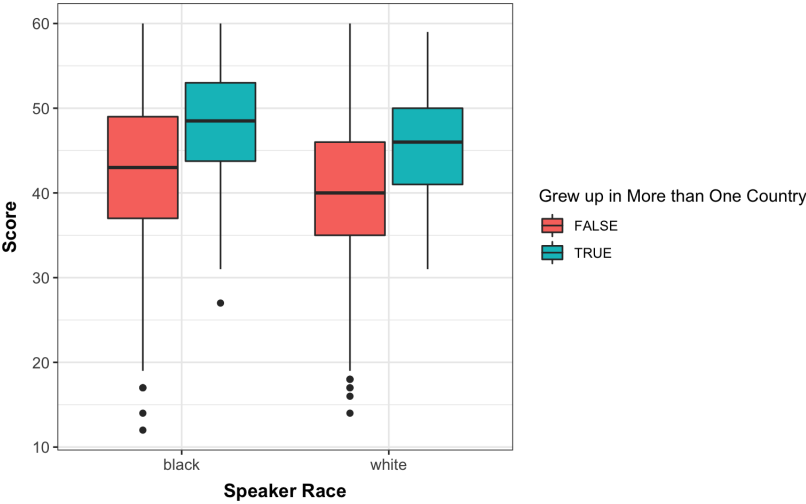FIGURE A.16. Scores by grown-up country and speaker race

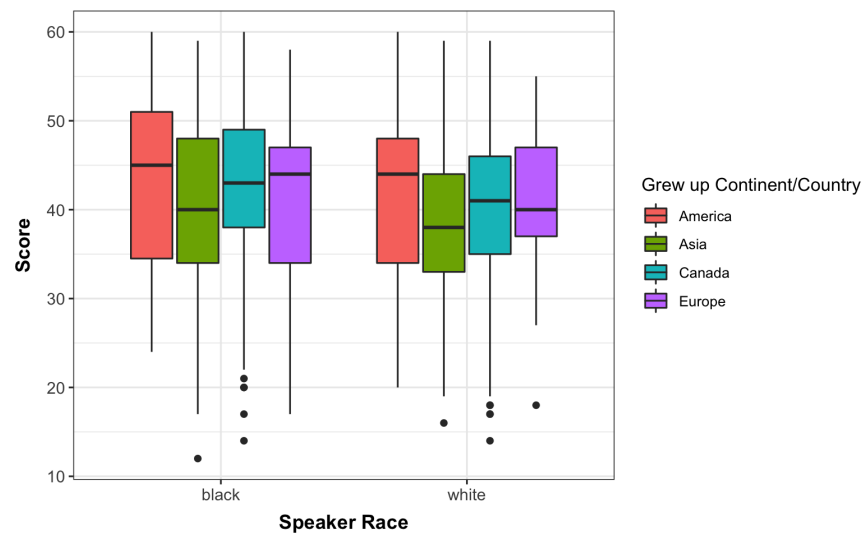FIGURE A.17. Scores by the number of grown-up countries and speaker race

FIGURE A.18. Scores by grown-up country/continent and speaker race