**Appendix B: Statistical Checks**

Since it is theoretically possible that the simplicity of a particular leader’s language could vary greatly from speech to speech, while that of another leader remains largely the same, we have checked whether

the results reported in our text (which apply to entire subcorpora) are consistent across each subcorpus. To do so, we have analysed the readability score distributions for each leader. Individual text readability values were calculated using the following websites:

* US and UK (Flesch-Kinkaid Reading Ease score): www.online-utility.org/english/readability\_test\_and\_improve.jsp
* France (Kandel-Moles readability index): visual-seo.com/it/SEO-Software-Features/Readability-Analysis
* Italy (Gulpease readability index): www.corrige.it.

The following tables report all speech readability scores for each subcorpus:

**Table B.1: Flesch-Kinkaid reading ease scores for each speech by US and UK leaders**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Speech No.** | **Trump** | **Clinton** | **Farage** | **Miliband** | **Cameron** |
| **1** | 62.11 | 60.93 | 54.92 | 65.88 | 69.21 |
| **2** | 74.43 | 70.59 | 62.56 | 65.72 | 70.75 |
| **3** | 59.7 | 63.19 | 49.56 | 63.97 | 74.98 |
| **4** | 75.78 | 70.46 | 58.77 | 67.04 | 73.3 |
| **5** | 77.62 | 70.94 | 60.39 | 75.72 | 58.74 |
| **6** | 55.94 | 73.62 | 59.91 | 54.5 | 69.34 |
| **7** | 67.18 | 71.56 | 57.05 | 76.92 | 62.2 |
| **8** | 62.15 | 62.05 | 60.83 | 71.45 | 70.67 |
| **9** | 62.6 | 71.17 | 63.82 | 66.01 | 67.3 |
| **10** | 58.35 | 70.66 | 59.59 | 70.4 | 60.19 |
| **11** | 56.58 | 72.51 | 61.34 | 76.69 | 59.4 |
| **12** | 55.41 | 66.61 | 57.3 | 72.46 | 68.8 |
| **13** | 76.7 | 64.61 | 52.86 | 71.83 | 72.6 |
| **14** | 58.77 | 66.25 | 53.09 | 69.92 | 61.59 |
| **15** | 72.66 | 72.32 | 59.09 | 73.22 | 72.41 |
| **16** | 75.73 | 71.3 | 63.39 | 63.62 | 70.77 |
| **17** | 64.69 | 75.24 | 62.27 | 71.99 | 65.53 |
| **18** | 63.39 | 66.09 | 65.15 | 75.71 | 70.4 |
| **19** | 65.73 | 60.15 | 59.17 | 73.33 | 68.11 |
| **20** | 66.94 | 59.48 | 60.85 | 71.88 | 61.09 |
| **21** | 72.16 | 71.34 | 61.54 | 66.95 | 63.22 |
| **22** | 79.33 | 63.33 | 62.54 | 65.38 | 63.5 |
| **23** |  | 64.93 | 61.67 | 69.54 | 53.71 |
| **24** |  | 65.58 | 62.32 | 69.35 | 65.39 |
| **25** |  | 67.26 | 68.21 | 73.17 | 68.15 |
| **26** |  | 65.27 | 50.96 | 70.81 | 70.52 |
| **27** |  | 71.54 | 60.6 | 62.85 |  |
| **28** |  | 70.05 | 64.35 | 67.22 |  |
| **29** |  | 69.32 | 55.86 | 62.37 |  |
| **30** |  | 62.17 | 43.03 | 62.29 |  |
| **31** |  |  | 54.16 | 62.63 |  |
| **32** |  |  | 51.1 |  |  |
| **33** |  |  | 60.94 |  |  |
| **34** |  |  | 57.66 |  |  |
| **35** |  |  | 54.08 |  |  |

**Table B.2: Kandel-Moles (France) and Gulpease (Italy) readability index scores for each speech by French and Italian leaders**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Speech No.** | **Le Pen** | **Macron** | **Salvini** | **Renzi** | **Alfano** |
| **1** | 52 | 65 | 61 | 53 | 48 |
| **2** | 60 | 61 | 61 | 58 | 51 |
| **3** | 55 | 67 | 56 | 57 | 48 |
| **4** | 60 | 68 | 64 | 64 | 52 |
| **5** | 61 | 60 | 65 | 59 | 51 |
| **6** | 61 | 65 | 61 | 58 | 52 |
| **7** | 60 | 65 | 65 | 58 | 51 |
| **8** | 55 | 61 | 68 | 58 | 52 |
| **9** | 58 | 61 | 65 | 51 | 51 |
| **10** | 60 | 70 | 63 | 55 | 54 |
| **11** | 57 | 62 | 58 | 59 | 49 |
| **12** | 61 | 67 | 58 | 55 | 49 |
| **13** | 62 | 63 | 62 | 55 | 51 |
| **14** | 49 | 65 | 64 | 66 | 46 |
| **15** | 39 | 63 | 60 | 52 | 47 |
| **16** | 51 | 68 | 66 | 59 | 48 |
| **17** | 56 |  | 54 | 52 | 48 |
| **18** | 54 |  | 60 | 51 | 51 |
| **19** | 51 |  | 63 | 54 | 52 |
| **20** | 56 |  | 59 | 54 | 48 |
| **21** | 52 |  | 61 |  | 48 |
| **22** |  |  | 61 |  | 48 |
| **23** |  |  | 75 |  | 49 |
| **24** |  |  | 56 |  | 55 |
| **25** |  |  | 61 |  |  |

We used the following analytical methods:

* graphical inspection of empirical density distribution (readability score histograms and empirical density distribution as in Figures B.1, B.2, B.3 and B.4).
* Shapiro-Wilks normality tests to evaluate distribution normality.
* F-test to evaluate distribution deviation differences.
* T‑test to evaluate mean distribution differences.

Since the T-test can only be applied when the distributions involved are normal, we used a Shapiro-Wilks normality test to ascertain distribution normality. The T‑test is sensitive to distribution variation, so distribution deviation was evaluated to parametrise the T‑test effectively. As the T-test was not applicable to one of the leaders, the Wilcoxon test was used to compare medians instead. Given the limited number of comparisons, we did not use ANOVA tests.

**Figure B.1: Flesch-Kincaid reading ease score distribution (United States)**



**Figure B.2: Kandel Moles readability index distribution (France)**

Since the Le Pen subcorpus included an outlier, two graphic representations (with and without the outlier) are presented below.



**Figure B.3: Flesch-Kincaid reading ease score distribution (United Kingdom)**



**Figure B.4: Gulpease readability index distribution (Italy)**



The Shapiro-Wilks normality test, F-test and T-test are evaluated by means of the p-value. The null hypotheses are reported below. The Shapiro-Wilks normality test was applied to the readability scores of all speeches within a political leader’s subcorpus. The null hypothesis is that the observed values belong to a normal distribution. P-values below 0.05 reject the null hypothesis. As shown in Table B.3 below, only Le Pen’s readability score does not pass the test because it includes a single outlier produced by a very low Kandel-Moles readability index. If that outlier is not considered in the sample, the null hypothesis is accepted.

**Table B.3: Shapiro-Wilks Normality test results**

|  |  |  |
| --- | --- | --- |
| **Leader** | **p-value** | **Normally distributed** |
| **Trump** | 0.112 | Yes |
| **Clinton** | 0.110 | Yes |
| **Le Pen** | 0.007 | No |
| **Le Pen - no outlier** | 0.101 | Yes |
| **Macron** | 0.405 | Yes |
| **Farage** | 0.072 | Yes |
| **Miliband** | 0.275 | Yes |
| **Cameron** | 0.258 | Yes |
| **Salvini** | 0.138 | Yes |
| **Renzi** | 0.127 | Yes |
| **Alfano** | 0.098 | Yes |

Having ascertained that the readability score distributions are normal, the T-test can be applied and parametrized to account for variance diversity among distributions. The table below reports two p-values: the first is to assess variance diversity, the second mean diversity. The F-test null hypothesis is that the variances in the distributions are the same. The T-test null hypothesis is that the means in the distributions are the same, with an alternative hypothesis that the populist leaders’ readability score is greater than that of the mainstream leader.

**Table B.4: F-test and T-test results**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Comparison** | **F-test p‑value** | **F-test p-value (explained)** | **T-test p‑value** | **T-test p‑value (explained)** |
| **Trump - Clinton** | 0.00 | Different variance | 0.73 | No difference in the mean |
| **Le Pen - Macron** | -- | -- | 1.00\* | Different means (Populist not simpler)\* |
| **Le Pen (without outlier) - Macron** | 0.27 | Same variance | 1.00 | Different means (Populist not simpler) |
| **Farage - Miliband** | 0.97 | Same variance | 1.00 | Different means (Populist not simpler) |
| **Farage - Cameron** | 0.83 | Same variance | 1.00 | Different means (Populist not simpler) |
| **Salvini - Renzi** | 0.75 | Same variance | 0.00 | Different means (Populist simpler) |
| **Salvini - Alfano** | 0.00 | Different variance | 0.00 | Different means (Populist simpler) |

\* As the distribution is not normal, the Wilcoxon test is used.