**SUPPORTING INFORMATION-A: SPEAKER QUESTIONNAIRE**

**Basic info**

1. Age: years old
2. Age of Arrival: years old
3. Where? ① UK ② North America ③ Australia/NZ ④ Others ( )
4. Why? ① Study abroad ② Work abroad ③ Immigration ④ Others ( )
5. Have you ever taught English before? years (e.g., 0-10 years)
6. Have you taken any linguistics classes/training before? (0 = no; 1 = yes)

**Length of residence**

1. UK: years
2. North America: years
3. Australia/NZ: years
4. Others: years (0 = NO) (which countries? )

**L2 English Learning in Classroom Settings**

1. Age of learning in classroom settings: years old
2. Length of learning in classroom settings: years

**First Language**

1. First language (from birth/most dominant): Which language? ( )

**L2 English Use Profile: Average over the past 1-2 years**

Frequency at work/school (professional settings)

1. L1/most dominant: % (0-100%)
2. English (with fluent speakers): % (0-100%)
3. English (with non-fluent speakers): % (0-100%)

Frequency with friends (social settings)

1. L1/most dominant: % (0-100%)
2. English (with fluent speakers): % (0-100%)
3. English (with non-fluent speakers): % (0-100%)

Frequency at home

1. L1/most dominant % (0-100%)
2. English (with fluent speakers): % (0-100%)
3. English (with non-fluent speakers): % (0-100%)

**Motivation and Concern for “Nativelikeness” (RP)**

***To what degree are you expected to use L2 English at a nativelike proficiency level on a 9-point scale (1 = not at all, 9 = very much so)?***

1. At work/school (professional settings)
2. With friends (social settings)
3. At home

**Awareness of one’s own L2 English oral proficiency**

***While speaking L2 English, which aspects of language do you think are relatively crucial for successful communication? Please rate the following statements on a 9-point scale (1 = not important, 9 = very important)?***

1. Speaking English without any accent like a native speaker
2. Speaking comprehensible English regardless of accentedness
3. Good pronunciation
4. Appropriate vocabulary/grammar
5. Idiomatic & sophisticated expression

**Perception of Foreign Accents**

***How much do you agree with the following statements (1 = strongly disagree, 9 = strongly agree)?***

1. I like it when people speak English with a foreign accent
2. I like it when people speak English with British accent (RP)
3. I like it when people speak English with American accent (GA).

***How much are you familiar with different types of English (1 = I am not familiar at all, 9 = I am very much)***

1. I am familiar with British English (Received Pronunciation) (1-9)
2. I am familiar with different kinds of foreign accented English (1-9)

**SUPPORTING INFORMATION-B: Descriptive Statistics of 30 Learner Variables Among 110 L2 Participants**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *M* | *SD* | *Range* | |
| *Min* | *Max* |
| First language |  |  |  |  |
| L1-L2 distance | Indo-European (46), Non Indo-European (64) | | | |
| A. Age |  |  |  |  |
| Chronological age | 30.4 | 7.4 | 20 | 59 |
| Age of arrival | 24.4 | 6.0 | 16 | 55 |
| Age of foreign language learning | 9.9 | 5.8 | 2 | 58 |
| B. Previous Experience |  |  |  |  |
| Length of residence in English speaking environments | 4.9 | 6.2 | 0.4 | 39 |
| Length of residence in the UK | 4.4 | 6.0 | 0.1 | 39 |
| Length of foreign language learning | 11.5 | 4.8 | 0 | 23 |
| Previous linguistics training experience | Yes (33), No (77) | | | |
| Previous English teaching experience | Yes (31), No (79) | | | |
| C. Current Experience |  |  |  |  |
| L1 use at work | 15.1% | 22.1 | 0 | 80 |
| L2 use with fluent speakers at work | 68.1% | 29.4 | 10 | 100 |
| L2 use with non-fluent speakers at work | 15.0% | 19.8 | 0 | 88 |
| L1 use in social settings | 40.2% | 28.7 | 0 | 100 |
| L2 use with fluent speakers in social settings | 48.4% | 29.2 | 0 | 100 |
| L2 use with non-fluent speakers in social settings | 8.4% | 14.2 | 0 | 80 |
| L1 use at home | 60.4% | 39.3 | 0 | 100 |
| L2 use with fluent speakers at home | 32.9% | 38.5 | 0 | 100 |
| L2 use with non-fluent speakers at home | 2.8% | 8.9 | 0 | 60 |
| D. Motivation |  |  |  |  |
| Expectation at work | 6.1 | 2.3 | 1 | 9 |
| Expectation in social settings | 4.6 | 2.4 | 1 | 9 |
| Expectation at home | 3.0 | 2.5 | 1 | 9 |
| E. Awareness |  |  |  |  |
| Awareness of nativelikeness | 5.0 | 2.3 | 1 | 9 |
| Awareness of comprehensibility | 8.2 | 1.0 | 4 | 9 |
| Awareness of pronunciation accuracy | 7.3 | 1.5 | 2 | 9 |
| Awareness of appropriate lexicogrammar | 6.9 | 1.7 | 1 | 9 |
| Awareness of idiomatic expression | 4.6 | 2.2 | 1 | 9 |
| F. Familiarity and Attitude |  |  |  |  |
| Familiarity towards foreign accent | 5.0 | 2.3 | 1 | 9 |
| Familiarity towards British English | 8.2 | 1.0 | 4 | 9 |
| Attitude towards foreign accent | 7.3 | 1.5 | 2 | 9 |
| Attitude towards British English | 6.9 | 1.7 | 1 | 9 |

**SUPPORTING INFORMATION-C: Training Scripts and Onscreen Labels of L2 Comprehensibility and Nativelikeness Judgements**

**A. Training scripts for comprehensibility and nativelikeness judgement**

|  |  |
| --- | --- |
| Comprehensibility | This term refers to how much effort it takes to understand what someone is saying. If you can understand with ease, then a speaker is highly comprehensible. However, if you struggle and must listen very carefully, or in fact cannot understand what is being said at all, then a speaker has low comprehensibility. |
| Nativelikeness | This refers to how much a second language speech sample differs from the variety of English commonly used in the community. |

**B. Onscreen labels**

|  |
| --- |
| **Comprehensibility**    1 2 3 4 5 6 7 8 9  Hard to understand Easy to understand |
| **Nativelikeness**  1 2 3 4 5 6 7 8 9  Not nativelike at all Very nativelike |

**SUPPORTING INFORMATION-D: Summary of a Ten-Factor Solution Based on a Factor Analysis of the Learner Background Questionnaire**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 | Factor 6 | Factor 7 | Factor 8 | Factor 9 | Factor 10 | Factor 11 |  |
| Labels | Experience Quantity | Current L2 Use | Awareness  of Nativeness | Age of Immersion | Motivation | Attitude to  Nativeness | EFL Experience | Special Past Experience | Foreign Accents Attitude | Comprehensibility Orientation | L1 Influence |  |
| Cumulative % | 15.39 | 27.14 | 36.82 | 46.07 | 54.07 | 59.88 | 65.29 | 70.03 | 74.59 | 78.76 | 82.55 |  |
| First language |  |  |  |  |  |  |  |  |  |  |  |  |
| L1-L2 distance | -0.019 | -0.090 | 0.135 | -0.018 | 0.032 | -0.092 | 0.110 | -0.055 | 0.180 | -0.015 | **0.902** |  |
| A. Age |  |  |  |  |  |  |  |  |  |  |  |  |
| Chronological age | **0.610** | 0.027 | 0.010 | **0.685** | -0.003 | -0.002 | -0.132 | -0.030 | 0.054 | 0.001 | -0.046 |  |
| Age of arrival | -0.186 | -0.044 | -0.009 | **0.924** | -0.032 | 0.057 | -0.026 | -0.016 | -0.065 | -0.012 | -0.025 |  |
| Age of foreign language learning | -0.048 | 0.014 | 0.073 | 0.264 | -0.037 | -0.051 | **-0.851** | -0.030 | -0.056 | 0.034 | 0.020 |  |
| B. Previous Experience |  |  |  |  |  |  |  |  |  |  |  |  |
| Length of residence in English speaking environments | **0.973** | 0.015 | -0.010 | -0.091 | 0.019 | -0.016 | -0.041 | -0.009 | 0.019 | -0.027 | 0.013 |  |
| Length of residence in the UK | **0.965** | 0.017 | -0.010 | -0.037 | 0.016 | -0.008 | -0.019 | 0.027 | 0.013 | -0.058 | -0.010 |  |
| Length of foreign language learning | -0.160 | 0.129 | 0.000 | 0.221 | -0.097 | -0.019 | **0.836** | 0.010 | -0.060 | 0.023 | 0.120 |  |
| Previous linguistics training experience | -0.069 | 0.042 | 0.055 | -0.041 | 0.030 | -0.012 | -0.104 | **-0.893** | 0.079 | -0.046 | 0.094 |  |
| Previous English teaching experience | 0.070 | 0.058 | -0.099 | 0.022 | -0.062 | 0.050 | 0.054 | **-0.926** | -0.100 | 0.090 | -0.047 |  |
| C. Current Experience |  |  |  |  |  |  |  |  |  |  |  |  |
| L1 use | 0.017 | **-0.901** | 0.019 | -0.016 | 0.032 | -0.040 | -0.046 | 0.066 | -0.057 | -0.058 | 0.026 |  |
| L2 use with fluent speakers | 0.035 | **0.928** | 0.081 | -0.031 | 0.008 | -0.040 | 0.029 | -0.049 | -0.040 | -0.001 | -0.021 |  |
| D. Motivation |  |  |  |  |  |  |  |  |  |  |  |  |
| Expectation at work | 0.000 | -0.124 | -0.026 | -0.084 | **0.825** | 0.078 | 0.038 | -0.039 | -0.084 | 0.277 | -0.063 |  |
| Expectation in social settings | 0.031 | -0.011 | 0.024 | -0.059 | **0.889** | 0.044 | -0.093 | 0.077 | -0.035 | 0.029 | 0.110 |  |
| Expectation at home | 0.011 | 0.157 | -0.016 | 0.184 | **0.649** | -0.033 | 0.037 | -0.062 | 0.193 | -0.376 | -0.058 |  |
| E. Awareness |  |  |  |  |  |  |  |  |  |  |  |  |
| Awareness of nativelikeness | -0.104 | 0.026 | **0.874** | -0.100 | -0.044 | 0.166 | -0.096 | -0.038 | -0.062 | -0.078 | 0.078 |  |
| Awareness of comprehensibility | -0.125 | 0.020 | 0.067 | 0.046 | 0.166 | -0.217 | 0.068 | -0.126 | -0.032 | **0.760** | 0.035 |  |
| Awareness of pronunciation accuracy | 0.167 | 0.022 | **0.668** | 0.171 | 0.080 | 0.084 | 0.087 | 0.075 | -0.134 | 0.239 | 0.158 |  |
| Awareness of appropriate lexicogrammar | 0.085 | -0.232 | 0.351 | 0.129 | -0.004 | 0.058 | 0.164 | -0.174 | 0.373 | 0.088 | -0.435 |  |
| Awareness of idiomatic expression | 0.008 | 0.070 | **0.610** | -0.001 | 0.054 | -0.282 | -0.025 | 0.019 | 0.184 | -0.068 | -0.426 |  |
| F. Familiarity and Attitude |  |  |  |  |  |  |  |  |  |  |  |  |
| Familiarity towards foreign accent | 0.007 | 0.191 | -0.051 | -0.087 | -0.077 | 0.264 | -0.092 | 0.143 | 0.318 | **0.605** | -0.150 |  |
| Familiarity towards British English | 0.180 | 0.099 | 0.110 | -0.196 | 0.086 | **0.721** | 0.173 | -0.046 | 0.022 | -0.015 | -0.116 |  |
| Attitude towards foreign accent | 0.041 | 0.050 | -0.099 | -0.039 | 0.000 | 0.008 | 0.000 | 0.009 | **0.927** | 0.031 | 0.161 |  |
| Attitude towards British English | -0.124 | -0.039 | 0.064 | 0.174 | 0.045 | **0.868** | -0.064 | 0.015 | 0.011 | -0.075 | 0.000 |  |

**SUPPORTING INFORMATION-E: Visual Inspection of Relationships Between L2 Speech Ratings and Learner Factors**

Below, we can see that there is some systematicity between factor scores and the ratings in some factors. Specifically, in both outcome measures, Factors 2, 6, and, to a lesser extent, 9 appear to be positively correlated with ratings while Factors 4, 8, and 11 appear to be negatively correlated. Variability of factor scores, however, tends to be large within each rating in each factor, and no firm conclusion can be drawn from the figure alone. We, therefore, tested the relationship in a more formal manner.



Figure 1. Distribution of factor scores in each rating of comprehensibility and nativelikeness in each factor. Grey dots represent individual judgments, while a larger dot represents the mean value in each rating category.

**SUPPORTING INFORMATION-G: Posterior Distribution**

Figures 1 through 3 show the posterior distribution of the population-level parameters, SDs of random intercepts, and correlation parameters in the Bayesian multivariate mixed-effects ordinal regression model. Most of the distributions appear to be reasonably normal, and even when they deviate from a normal distribution, the CIs seem summarise the distribution to a good extent (e.g., SDs of random intercepts in Figure 3). The only possible exception is the correlation of by-participant random intercepts (i.e., lower-mid panel in Figure 3), which appears to have reached the ceiling and the mean of the distribution differs from its mode. The interpretation of the parameter, therefore, requires some caution. We do not interpret it.

****

Figure 1. Posterior distribution of the fixed-effects parameters in comprehensibility ratings. Blue horizontal lines represent 95% CIs, while red vertical lines (only drawn for slope parameters) represent the null effect.

****

Figure 2. Posterior distribution of the fixed-effects parameters in nativelikeness ratings. See the caption of Figure 1 for the interpretation.

****

Figure 3. Posterior distribution of random intercepts and correlation parameters. See the caption of Figure 1 for the interpretation.

**SUPPORTING INFORMATION-H: Sensitivity Analysis**

Figures 1 and 2 show the posterior means and their 95% CIs across different prior distributions in comprehensibility ratings and nativelikeness ratings, respectively. We can observe that neither the point estimate of the parameter nor its uncertainty is affected much by the choice of the priors considered here. The inferences drawn, therefore, is considered fairly robust against the choice of priors.



Figure 1. Posterior means and their 95% CIs in each slope parameter across different choices of prior distribution in comprehensibility ratings. SD represents the SD of a normal distribution, while flat corresponds to a flat prior



Figure 2. Posterior means and their 95% CIs in each slope parameter across different choices of prior distribution in nativelikeness ratings.

**SUPPORTING INFORMATION-I: Correlates of L2 Comprehensibility vs. Nativelikeness at Different Ability Levels**

To grasp the transition of the importance of each factor across the rating scale, for each outcome variable, a series of eight Bayesian mixed-effects binary logistic regression models were fitted to the subset of the data that each only included observations with two adjacent ratings (e.g., 1 vs 2, 2 vs 3, 3 vs 4). The purpose here is to see if and how the posterior mean and its CIs representing the strength of predictors change between low-rating contrasts (e.g., 1 vs 2) and high-rating contrasts (e.g., 8 vs 9). The fixed-effects component of the models included the four (in comprehensibility) or five (in nativelikeness) factors whose 95% CIs did not include 0 in the ordinal regression model, and the random-effects component included by-learner and by-rater random intercepts. The specification of prior distribution is similar to the model we have discussed above, except that flat priors were used for slope parameters.

Figure 1 shows the posterior mean and the 95% CIs of the predictors across the eight models in each outcome measure. The contrast between the two lowest ratings in comprehensibility (i.e., 1 vs 2) is not presented due to the rather small sample size and absurdly large CIs as its results. In the figure, it is, for instance, observed that the effect of Factor 2 (Current L2 Use) is positive in the contrast between 2 and 3 in comprehensibility, which indicates that, in the model that only targeted speech samples rated as 2 or 3, the probability that a sample is given the rating of 3 increases as the factor scores of Factor 2 increase. The effect, however, is trivial, as the 95% credible interval crosses 0. Indeed, most of the models in the figure included 0 in their 95% CIs of the slope parameter, presumably due to the small sample size included in each model. This, however, is not a major issue, as the purpose here is to identify potentially interesting transitional patterns of posterior means.

When each panel of comprehensibility is examined vertically, Factors 2 (Current L2 Use), 4 (Age of Immersion), and 8 (Special Past Experience) do not demonstrate clear patterns, as their posterior means appear to fluctuate randomly from one model to another. In Factor 6 (Attitude to Nativeness), however, there appears to be an increasing trend in that the expected effect of Factor 6 starts at around 0 (i.e., factor scores are unrelated to ratings) and ends positively (i.e., higher scores are associated with higher ratings). Attitude to nativeness, therefore, may influence the discrimination of comprehensibility only at higher ratings. Similarly, regarding nativelikeness ratings, Factors 4 (Age of Immersion) and 11 (L1 Influence) appear to exert stronger influence at higher ratings. The figure, therefore, suggests that the strength of some factors can well vary across the scale of comprehensibility and nativelikeness ratings.



*Figure 1*. The so-called ‘secret weapon’ figure (Gelman & Hill, 2007) which shows the posterior means and their 95% CIs in each of the eight Bayesian mixed-effects logistic regression models targeting the observations with two adjacent ratings in each outcome measure.

***Reference***

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York, NY: Cambridge University Press.