**Supporting information S1**: Preliminary analysis for the speech samples

An R code for the Cluster analysis (N = 102):

install.packages("plotly")

install.packages("ggdendro")

install.packages("cluster")

library(cluster)

library(plotly)

library(ggdendro)

# Hierarchical clustering analysis for N = 102 speech rate in phone topic

SRph <- read.csv("speechrate\_phone.csv", row.names=1)

head(SRph)

SRphdist <- dist(SRph,method="euclidean") #determine distance

SRphhc <- hclust(SRphdist, "ward.D2") #cluster analysis

plot(SRphhc) #plot dendrogram

# Compute k-means

km.resph <- kmeans(SRph, 2, nstart = 25) #divide sample into 2 clusters

head(km.resph$cluster, 102) #view dataset labeled either 1 or 2

silph <- silhouette(km.resph$cluster, dist(SRph)) #silhouette analysis

rownames(silph) <- rownames(SRph)

head(silph[, 1:3])

fviz\_silhouette(silph) #plot silhouette result

A diagram of a cluster dendrogram

Description automatically generated A graph showing different colored squares

Description automatically generated with medium confidence

**Figure S1**. Dendrogram plot and Silhouette plot

**Supporting information S2:** Descriptive statistics of PF, UF, and ngram measures

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S2.1**. Descriptive Statistics of PF and UF measures | | | | | | | |
|  |  |  | 95% *CI* (*M*) | | Shapiro-Wilk *p*-value |  |  |
|  | *M* | *SD* | lower | upper | Min | Max |
| Perceived fluency | 3.71 | 1.28 | 3.46 | 3.96 | 0.036 | 1 | 6 |
| *Speed fluency* |  |  |  |  |  |  |  |
| Articulation rate | 2.83 | 0.68 | 2.70 | 2.96 | 0.181 | 1.28 | 4.22 |
| *Breakdown fluency* |  |  |  |  |  |  |  |
| Midclause pause ratio | 0.21 | 0.13 | 0.19 | 0.24 | < .001 | 0.03 | 0.67 |
| Endclause pause ratio | 0.07 | 0.03 | 0.07 | 0.08 | 0.253 | 0.02 | 0.14 |
| Midclause pause duration (sec) | 0.84 | 0.30 | 0.78 | 0.90 | < .001 | 0.45 | 1.73 |
| Endclause pause duration (sec) | 0.88 | 0.35 | 0.81 | 0.95 | < .001 | 0.37 | 2.25 |
| Filled pause ratio | 0.09 | 0.09 | 0.07 | 0.11 | < .001 | 0 | 0.47 |
| *Repair fluency* |  |  |  |  |  |  |  |
| Disfluency ratio | 0.11 | 0.09 | 0.09 | 0.13 | < .001 | 0 | 0.50 |
| *Note*. *N* = 102 | | | | | | | |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S2.2**. Descriptive statistics of ngram measures | | | | | | | |
|  |  |  | 95% *CI* (*M*) | | Shapiro-Wilk *p*-value |  |  |
|  | *M* | *SD* | lower | upper | Min | Max |
| *Bigram* |  |  |  |  |  |  |  |
| Frequency (log)a | 1.19 | 0.18 | 1.16 | 1.23 | 0.286 | 0.73 | 1.55 |
| Mutual informationa | 1.14 | 0.16 | 1.11 | 1.17 | 0.032 | 0.69 | 1.71 |
| Proportion (30k) | 0.59 | 0.08 | 0.58 | 0.61 | 0.689 | 0.39 | 0.77 |
| *Trigram* |  |  |  |  |  |  |  |
| Frequency (log)a | 0.33 | 0.11 | 0.31 | 0.35 | 0.724 | 0.07 | 0.64 |
| Mutual information (uni→bi)a | 0.75 | 0.19 | 0.71 | 0.79 | 0.691 | 0.20 | 1.24 |
| Mutual information (bi→uni)a | 0.71 | 0.20 | 0.68 | 0.75 | 0.64 | 0.21 | 1.14 |
| Proportion (30k) | 0.23 | 0.07 | 0.22 | 0.25 | 0.806 | 0.06 | 0.41 |
| aSum of the indice scores were divided by text length | | | | | | | |

**Supporting information S3:** Correlations between (1) PF and UF and (2) UF and ngram indices (frequency, MI, and proportion)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table S3.1**. Spearman's correlations between PF and UF measures | | | | | |
|  | Perceived fluency | | | | |
|  |  |  |  | 95% CI | |
|  | *rho* |  | *p* | lower | upper |
| *Speed fluency* |  |  |  |  |  |
| Articulation rate | 0.60 | \*\*\* | < .001 | 0.46 | 0.71 |
| *Breakdown fluency* |  |  |  |  |  |
| Midclause pause ratio | -0.84 | \*\*\* | < .001 | -0.89 | -0.78 |
| Endclause pause ratio | -0.66 | \*\*\* | < .001 | -0.76 | -0.54 |
| Midclause pause duration (sec) | -0.81 | \*\*\* | < .001 | -0.87 | -0.73 |
| Endclause pause duration (sec) | -0.64 | \*\*\* | < .001 | -0.75 | -0.51 |
| Filled pause ratio | -0.18 |  | 0.07 | -0.36 | 0.02 |
| *Repair fluency* |  |  |  |  |  |
| Disfluency ratio | -0.49 | \*\*\* | < .001 | -0.63 | -0.33 |
| *Note*. \* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001 | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S3.2**. Spearman's correlations between UF and bigram measures | | | | | | | | | | | | | | |
|  | Bigram | | | | | | | | | | | | | |
|  | Frequencya | | | |  | MIa | | | |  | Proportion 30k | | | |
|  |  |  | 95% CI | |  |  |  | 95% CI | |  |  |  | 95% CI | |
|  | *rho* | | lower | upper |  | *rho* | | lower | upper |  | *rho* | | lower | upper |
| *Speed fluency* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Articulation rate | 0.26 | \*\* | 0.01 | 0.07 |  | 0.31 | \*\* | -0.11 | 0.27 |  | 0.17 |  | -0.03 | 0.35 |
| *Breakdown fluency* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Midclause pause ratio | -0.28 | \*\* | 0.01 | -0.45 |  | -0.31 | \*\* | -0.22 | 0.17 |  | -0.26 | \*\* | -0.43 | -0.07 |
| Endclause pause ratio | -0.05 |  | 0.64 | -0.24 |  | 0.00 |  | -0.10 | 0.28 |  | -0.05 |  | -0.24 | 0.15 |
| Midclause pause duration (sec) | -0.25 | \* | 0.01 | -0.43 |  | -0.03 |  | 0.01 | 0.38 |  | -0.23 | \* | -0.41 | -0.04 |
| Endclause pause duration (sec) | -0.17 |  | 0.09 | -0.35 |  | 0.01 |  | -0.01 | 0.37 |  | -0.20 | \* | -0.38 | 0.00 |
| Filled pause ratio | 0.01 |  | 0.89 | -0.18 |  | -0.27 | \*\* | -0.41 | -0.04 |  | 0.09 |  | -0.11 | 0.28 |
| *Repair fluency* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Disfluency ratio | -0.17 |  | 0.09 | -0.35 |  | -0.33 | \*\*\* | -0.34 | 0.04 |  | -0.07 |  | -0.26 | 0.13 |
| *Note*. aSum of the indice scores were divided by text length; MI = mutual information; \* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001 | | | | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S3.3**. Spearman's correlations between UF and trigram measures | | | | | | | | | | | | | | | |
|  | Trigram | | | | | | | | | | | | | | |
|  | Frequencya | | | |  | MI (uni→bi)a | | | |  | Proportion 30k | | | |
|  |  |  | 95% CI | |  |  |  | 95% CI | |  |  |  | 95% CI | |
|  | *rho* | | lower | upper |  | *rho* | | lower | upper |  | *rho* | | lower | upper |
| *Speed fluency* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Articulation rate | 0.25 | \* | 0.01 | 0.06 |  | 0.27 | \*\* | 0.08 | 0.44 |  | 0.29 | \*\* | 0.10 | 0.46 |
| *Breakdown fluency* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Midclause pause ratio | -0.22 | \* | 0.03 | -0.40 |  | -0.33 | \*\*\* | -0.49 | -0.14 |  | -0.25 | \* | -0.43 | -0.06 |
| Endclause pause ratio | -0.04 |  | 0.66 | -0.24 |  | -0.09 |  | -0.28 | 0.11 |  | -0.05 |  | -0.24 | 0.15 |
| Midclause pause duration (sec) | -0.18 |  | 0.07 | -0.36 |  | -0.28 | \*\* | -0.45 | -0.09 |  | -0.23 | \* | -0.40 | -0.03 |
| Endclause pause duration (sec) | -0.15 |  | 0.14 | -0.33 |  | -0.23 | \* | -0.41 | -0.04 |  | -0.14 |  | -0.32 | 0.06 |
| Filled pause ratio | 0.02 |  | 0.85 | -0.18 |  | -0.03 |  | -0.22 | 0.17 |  | 0.02 |  | -0.18 | 0.21 |
| *Repair fluency* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Disfluency ratio | -0.12 |  | 0.25 | -0.30 |  | -0.10 |  | -0.29 | 0.10 |  | -0.10 |  | -0.29 | 0.10 |
| *Note*. aSum of the indice scores were divided by text length; MI = mutual information; \* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001 | | | | | | | | | | | | | | | |

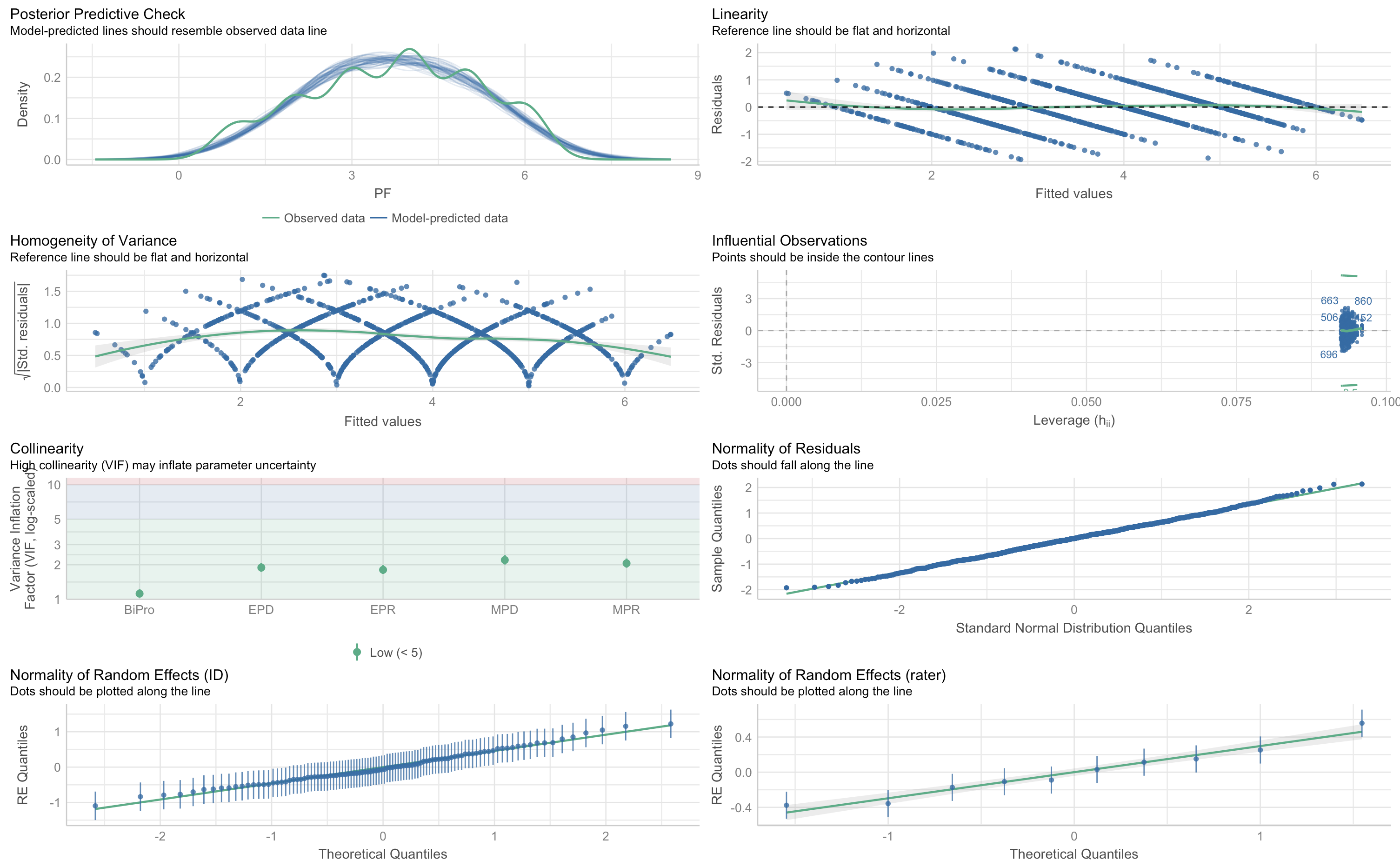
**Supporting information S4**: Partial correlations between ngram and PF while controlling for the effects of UF measure

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table S4.1**. Spearman correlations between PF and frequency with UF measures partialled out | | | | | | |
|  | Frequency (log) | | | | | |
|  | Bigram | | | Trigram | | |
| UF covariates | *rho* | | *p* | *rho* | | *p* |
| *Speed fluency* |  |  |  |  |  |  |
| Articulation rate | 0.21 | \* | 0.033 | 0.12 |  | 0.226 |
| *Breakdown fluency* |  |  |  |  |  |  |
| Midclause pause ratio | 0.18 |  | 0.078 | 0.11 |  | 0.269 |
| Endclause pause ratio | 0.39 | \*\*\* | < .001 | 0.29 | \*\* | 0.004 |
| Midclause pause duration (sec) | 0.21 | \* | 0.037 | 0.17 |  | 0.092 |
| Endclause pause duration (sec) | 0.29 | \*\* | 0.004 | 0.19 |  | 0.052 |
| Filled pause ratio | 0.33 | \*\*\* | < .001 | 0.25 | \* | 0.012 |
| *Repair fluency* |  |  |  |  |  |  |
| Disfluency ratio | 0.28 | \*\* | 0.005 | 0.22 | \* | 0.031 |
| \* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001 | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table S4.2**. Spearman correlations between PF and MI with UF measures partialled out | | | | | | |
|  | MI | | | | | |
|  | Bigram | | | Trigram (uni→bi) | | |
| UF covariates | *rho* | | *p* | *rho* | | *p* |
| *Speed fluency* |  |  |  |  |  |  |
| Articulation rate | -0.07 |  | 0.477 | 0.21 | \* | 0.039 |
| *Breakdown fluency* |  |  |  |  |  |  |
| Midclause pause ratio | -0.25 | \* | 0.012 | 0.09 |  | 0.383 |
| Endclause pause ratio | 0.18 |  | 0.076 | 0.36 | \*\*\* | < .001 |
| Midclause pause duration (sec) | 0.18 |  | 0.073 | 0.17 |  | 0.09 |
| Endclause pause duration (sec) | 0.18 |  | 0.075 | 0.24 | \* | 0.018 |
| Filled pause ratio | 0.09 |  | 0.387 | 0.32 | \*\* | 0.001 |
| *Repair fluency* |  |  |  |  |  |  |
| Disfluency ratio | -0.04 |  | 0.705 | 0.32 | \*\* | 0.001 |
| \* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001 | | | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table S4.3**. Spearman correlations between PF and proportion measures with UF measures partialled out | | | | | | |
|  | Proportion 30k | | | | | |
|  | Bigram | | | Trigram | | |
| UF covariates | *rho* | | *p* | *rho* | | *p* |
| *Speed fluency* |  |  |  |  |  |  |
| Articulation rate | 0.29 | \*\* | 0.003 | 0.11 |  | 0.286 |
| *Breakdown fluency* |  |  |  |  |  |  |
| Midclause pause ratio | 0.22 | \* | 0.029 | 0.09 |  | 0.39 |
| Endclause pause ratio | 0.40 | \*\*\* | < .001 | 0.30 | \*\* | 0.002 |
| Midclause pause duration (sec) | 0.25 | \* | 0.012 | 0.13 |  | 0.197 |
| Endclause pause duration (sec) | 0.27 | \*\* | 0.006 | 0.23 | \* | 0.024 |
| Filled pause ratio | 0.36 | \*\*\* | < .001 | 0.26 | \*\* | 0.008 |
| *Repair fluency* |  |  |  |  |  |  |
| Disfluency ratio | 0.34 | \*\*\* | < .001 | 0.24 | \* | 0.015 |
| \* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001 | | | | | | |

**Supporting information S5**: Assumption check of the final mixed-effect model by *performance* package (Lüdecke et al., 2021)



**Figure S5.1**. Model formula: PF ~ MPR + MPD + EPR + EPD + BiPro + (1 | ID) + (1 | rater)

**Supporting information S6**: Factors influencing rater perception of fluency

|  |  |
| --- | --- |
| **Table S6.1**. Factors involved in fluency perceptions | |
| Coded category | Count |
| Pronunciation | 23 |
| Prompt effect | 9 |
| Familiarity to Japanese | 10 |
| Speech duration | 6 |
| Content | 7 |
| UF-related difficulty | 4 |
| Grammar | 2 |
| Phrasal competence | 3 |
| L1 use | 3 |
| Strategy | 2 |
| L1 speaking style | 2 |
| Volume | 1 |
| Comprehensibility | 1 |
| Intelligibility | 1 |

**Supporting information S7**: Top and bottom five participants in PF rating scores

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table S7.1**. Top and bottom five participants for PF score with ngram indices scores | | | | | | | | | | | | | | | | |
| Rank | PF score | Text length | Bigram | | | | |  | Trigram | | | | | | |
| FREQa | FREQb | MIa | MIb | PROP |  | FREQa | FREQb | MIa | MIb | MI2a | MI2b | PROP |
| 1 | 6 | 246 | 1.60 | 1.25 | 1.41 | 1.10 | 0.65 |  | 0.92 | 0.31 | 2.46 | 0.83 | 2.34 | 0.79 | 0.22 |
| 2 | 6 | 124 | 1.37 | 0.94 | 2.09 | 1.43 | 0.51 |  | 0.98 | 0.23 | 3.34 | 0.78 | 2.97 | 0.69 | 0.15 |
| 3 | 6 | 123 | 1.73 | 1.39 | 1.46 | 1.17 | 0.64 |  | 1.13 | 0.41 | 2.15 | 0.77 | 1.99 | 0.71 | 0.25 |
| 4 | 5.9 | 158 | 1.86 | 1.34 | 1.54 | 1.11 | 0.60 |  | 1.18 | 0.42 | 2.06 | 0.73 | 2.15 | 0.76 | 0.27 |
| 5 | 5.9 | 211 | 1.76 | 1.44 | 1.42 | 1.16 | 0.70 |  | 1.13 | 0.49 | 2.24 | 0.98 | 2.25 | 0.98 | 0.32 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 98 | 1.6 | 103 | 1.66 | 1.21 | 1.84 | 1.34 | 0.57 |  | 1.12 | 0.31 | 2.50 | 0.68 | 2.59 | 0.70 | 0.20 |
| 99 | 1.5 | 64 | 1.56 | 1.07 | 1.80 | 1.24 | 0.52 |  | 1.42 | 0.33 | 2.82 | 0.66 | 2.61 | 0.61 | 0.19 |
| 100 | 1.4 | 68 | 1.69 | 1.14 | 1.52 | 1.03 | 0.58 |  | 1.01 | 0.31 | 2.26 | 0.70 | 2.29 | 0.71 | 0.24 |
| 101 | 1.3 | 26 | 1.83 | 1.20 | 1.50 | 0.98 | 0.60 |  | 1.21 | 0.33 | 1.97 | 0.53 | 1.74 | 0.47 | 0.23 |
| 102 | 1.1 | 102 | 1.46 | 0.76 | 1.32 | 0.69 | 0.39 |  | 0.74 | 0.07 | 2.22 | 0.20 | 2.43 | 0.21 | 0.06 |
| aMean-based indices were calculated by sum of the indices scores divided by total number of ngrams with the scores bNew measures were calculated by sum of the indices scores divided by text length, which is similar to the calculation in proportion index | | | | | | | | | | | | | | | | |

Following the reviewer’s suggestion, we conducted correlation analyses between (1) mean-based bigram/trigram frequency and MI and text length, and (2) text-length-controlled bigram/trigram frequency and MI and text length:

(1)

* Mean-based bigram frequency and text length (rho = .024, p = .810)
* Mean-based bigram MI and text length (rho = **-.261**, p = .008)
* Mean-based trigram frequency and text length (rho = **-.200**, p = .044)
* Mean-based trigram MI (unigram -> bigram) and text length (rho = **-.322**, p < .001)
* Mean-based trigram MI2 (bigram -> unigram) and text length (rho = **-.223**, p = .024)

(2)

* Text-length-controlled bigram frequency and text length (rho = **.210**, p = .034)
* Text-length-controlled bigram MI and text length (rho = .024, p = .807)
* Text-length-controlled trigram frequency and text length (rho = .100, p = .318)
* Text-length-controlled trigram MI (unigram -> bigram) and text length (rho = .156, p = .117)
* Text-length-controlled trigram MI2 (bigram -> unigram) and text length (rho = .181, p = .069)

For example, the top-ranked participant (PF score = 6) produced 246 words and their mean-based bigram and trigram MI are 1.41 and 2.46. In contrast, the participants ranked 99th (PF score = 1.5) produced 64 words and their mean-based bigram and trigram MI are 1.80 and 2.82.

As a reference, table below provides concrete trigrams and the MI scores for the two participants (rank 1st vs. rank 99th):

*Note*. Non-detected trigrams are omitted from the table for clarity.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Table S7.2**. Comparison between top-ranked participant and 99th ranked participant for the PF rating score | | | | | |
| Participant X | Trigram | MI (uni -> bi) | Participant Y | Trigram | MI (uni -> bi) |
| PF score = 6 (rank = 1st) | *first of all* | 5.894550924 | PF score = 1.5 (rank = 99th) | *many many time* | 4.133569542 |
| Mean-based MI score = 2.46 | *reason why i* | 5.658235791 | Mean-based MI score = 2.82 | *do not agree* | 4.079528388 |
| Text length = 246 words | *how much you* | 5.633268654 | Text length = 64 words | *do not agree* | 4.079528388 |
|  | *keep in touch* | 5.13282237 |  | *in order to* | 3.976098926 |
|  | *for example if* | 4.779096553 |  | *effect on we* | 3.850112241 |
|  | *for example when* | 4.683319876 |  | *effect on we* | 3.850112241 |
|  | *disagree with this* | 4.132966827 |  | *i do not* | 2.738084136 |
|  | *disagree with this* | 4.132966827 |  | *i do not* | 2.738084136 |
|  | *do not have* | 4.064580342 |  | *go to the* | 2.507893969 |
|  | *something to eat* | 3.955689796 |  | *not agree with* | 2.411477718 |
|  | *anyone who have* | 3.931876015 |  | *not agree with* | 2.411477718 |
|  | *effect on we* | 3.850112241 |  | *time in my* | 2.079720755 |
|  | *in touch with* | 3.844399608 |  | *order to go* | 1.649824837 |
|  | *of course have* | 3.812673718 |  | *so i do* | 1.570051084 |
|  | *reason for this* | 3.751999744 |  | *to go to* | 0.262130991 |
|  | *use the internet* | 3.73487297 |  |  |  |
|  | *contact with you* | 3.687808634 |  |  |  |
|  | *second of all* | 3.633875956 |  |  |  |
|  | *if you be* | 3.465078515 |  |  |  |
|  | *if you be* | 3.465078515 |  |  |  |
|  | *there be always* | 3.447263497 |  |  |  |
|  | *not have any* | 3.439209899 |  |  |  |
|  | *you can always* | 3.348371855 |  |  |  |
|  | *you can always* | 3.348371855 |  |  |  |
|  | *you can always* | 3.348371855 |  |  |  |
|  | *you can always* | 3.348371855 |  |  |  |
|  | *a century ago* | 3.220530678 |  |  |  |
|  | *you can use* | 3.19553667 |  |  |  |
|  | *you can use* | 3.19553667 |  |  |  |
|  | *want to find* | 3.130439412 |  |  |  |
|  | *you have question* | 3.106981277 |  |  |  |
|  | *when you be* | 3.011048769 |  |  |  |
|  | *whenever you have* | 2.863561023 |  |  |  |
|  | *whenever you have* | 2.863561023 |  |  |  |
|  | *with you friend* | 2.755946227 |  |  |  |
|  | *you can keep* | 2.733794831 |  |  |  |
|  | *of all i* | 2.7331814 |  |  |  |
|  | *in contact with* | 2.674406436 |  |  |  |
|  | *we daily life* | 2.651641438 |  |  |  |
|  | *which be really* | 2.582524572 |  |  |  |
|  | *be really amazing* | 2.578989819 |  |  |  |
|  | *be always a* | 2.569991057 |  |  |  |
|  | *ago when they* | 2.534013793 |  |  |  |
|  | *i disagree with* | 2.486073058 |  |  |  |
|  | *i disagree with* | 2.486073058 |  |  |  |
|  | *to find some* | 2.442319339 |  |  |  |
|  | *example if you* | 2.354534256 |  |  |  |
|  | *they have all* | 2.173448802 |  |  |  |
|  | *you have to* | 2.073479057 |  |  |  |
|  | *see how much* | 2.041018716 |  |  |  |
|  | *to eat you* | 1.950647036 |  |  |  |
|  | *you want to* | 1.938552057 |  |  |  |
|  | *and of course* | 1.934100454 |  |  |  |
|  | *so you can* | 1.920809596 |  |  |  |
|  | *or something to* | 1.9167203 |  |  |  |
|  | *to go or* | 1.879064109 |  |  |  |
|  | *they do not* | 1.872010084 |  |  |  |
|  | *these be the* | 1.869362418 |  |  |  |
|  | *when they do* | 1.863732853 |  |  |  |
|  | *go out you* | 1.848440122 |  |  |  |
|  | *in this way* | 1.81729887 |  |  |  |
|  | *can use the* | 1.622380579 |  |  |  |
|  | *so there be* | 1.551036882 |  |  |  |
|  | *you be go* | 1.460146875 |  |  |  |
|  | *so for example* | 1.357410162 |  |  |  |
|  | *have to find* | 1.344439852 |  |  |  |
|  | *that you can* | 1.332908301 |  |  |  |
|  | *the answer for* | 1.299981638 |  |  |  |
|  | *so i have* | 1.250591747 |  |  |  |
|  | *and these be* | 1.217733117 |  |  |  |
|  | *and the answer* | 1.034627804 |  |  |  |
|  | *or if you* | 1.029671571 |  |  |  |
|  | *of all of* | 0.651477041 |  |  |  |
|  | *have all those* | 0.525978901 |  |  |  |
|  | *answer for the* | 0.464649013 |  |  |  |
|  | *be go out* | 0.090597189 |  |  |  |
|  | *be the three* | 0.056965993 |  |  |  |
|  | *and outside of* | -0.006358475 |  |  |  |
|  | *or they have* | -0.176109543 |  |  |  |
|  | *so in this* | -0.3024575 |  |  |  |
|  | *for the question* | -0.541710079 |  |  |  |
|  | *and for example* | -1.029025633 |  |  |  |
|  | *all of course* | -1.141265375 |  |  |  |

*Note***.** Non-detected trigrams are omitted from the table for clarity.

**Supporting information S8**: Statistical analyses and R code

library(lme4)

library(lmerTest)

library(MuMIn)

library(car)

library(dplyr)

library(performance)

df <- read.csv("df\_long.csv",stringsAsFactors = F)

df\_select <- df %>%

dplyr::select(ID,PF,rater,ArticulationRate,Midclause.Silent.Pause.Ratio,Endclause.Silent.Pause.Ratio,

Midclause.Silent.Pause.Duration,Endclause.Silent.Pause.Duration,Filled.Pause.Ratio,DisfluencyRatio,COCA\_lemma\_spoken\_bi\_prop\_30k,COCA\_lemma\_spoken\_tri\_prop\_30k,bigram\_MI\_NEW,trigram\_MI\_NEW,bigram\_Fre\_NEW,trigram\_Fre\_NEW)

summary(df\_select)

#scaling and centering

AR <- scale(df\_select$ArticulationRate,center=TRUE,scale=TRUE)

summary(AR)

MPR <- scale(df\_select$Midclause.Silent.Pause.Ratio,center=TRUE,scale=TRUE)

summary(MPR)

EPR <- scale(df\_select$Endclause.Silent.Pause.Ratio,center=TRUE,scale=TRUE)

summary(EPR)

MPD <- scale(df\_select$Midclause.Silent.Pause.Duration,center=TRUE,scale=TRUE)

summary(MPD)

EPD <- scale(df\_select$Endclause.Silent.Pause.Duration,center=TRUE,scale=TRUE)

summary(EPD)

FPR <- scale(df\_select$Filled.Pause.Ratio,center=TRUE,scale=TRUE)

summary(FPR)

DR <- scale(df\_select$DisfluencyRatio,center=TRUE,scale=TRUE)

summary(DR)

BiF.N <- scale(df\_select$bigram\_Fre\_NEW,center=TRUE,scale=TRUE)

summary(BiF.N)

BiMI.N <- scale(df\_select$bigram\_MI\_NEW,center=TRUE,scale=TRUE)

summary(BiMI.N)

BiPro <- scale(df\_select$COCA\_lemma\_spoken\_bi\_prop\_30k,center=TRUE,scale=TRUE)

summary(BiPro)

TriF.N <- scale(df\_select$trigram\_Fre\_NEW,center=TRUE,scale=TRUE)

summary(TriF.N)

TriMI.N <- scale(df\_select$trigram\_MI\_NEW,center=TRUE,scale=TRUE)

summary(TriMI.N)

TriPro <- scale(df\_select$COCA\_lemma\_spoken\_tri\_prop\_30k,center=TRUE,scale=TRUE)

summary(TriPro)

### Modeling PF w/ most parsimonious UF variables ###

# unconditional model

LMM.PARS <- list()

LMM.PARS[[1]] <- lmer(PF~(1|ID)+(1|rater),data=df\_select,REML=FALSE,control=lmerControl(optimizer="bobyqa"))

# model 1

cand <- list()

cand[[1]] <- update(LMM.PARS[[1]], . ~ . + AR)

cand[[2]] <- update(LMM.PARS[[1]], . ~ . + MPR)

cand[[3]] <- update(LMM.PARS[[1]], . ~ . + EPR)

cand[[4]] <- update(LMM.PARS[[1]], . ~ . + MPD)

cand[[5]] <- update(LMM.PARS[[1]], . ~ . + EPD)

cand[[6]] <- update(LMM.PARS[[1]], . ~ . + FPR)

cand[[7]] <- update(LMM.PARS[[1]], . ~ . + DR)

sapply(cand, AIC) %>% data.frame

sapply(cand, AIC) %>% which.min

sapply(LMM.PARS, AIC)

LMM.PARS[[2]] <- cand[[2]] # MPR is added

# model 2

cand <- list()

cand[[1]] <- update(LMM.PARS[[2]], . ~ . + AR)

cand[[2]] <- update(LMM.PARS[[2]], . ~ . + EPR)

cand[[3]] <- update(LMM.PARS[[2]], . ~ . + MPD)

cand[[4]] <- update(LMM.PARS[[2]], . ~ . + EPD)

cand[[5]] <- update(LMM.PARS[[2]], . ~ . + FPR)

cand[[6]] <- update(LMM.PARS[[2]], . ~ . + DR)

sapply(cand, AIC) %>% data.frame

sapply(cand, AIC) %>% which.min

sapply(LMM.PARS, AIC)

LMM.PARS[[3]] <- cand[[3]] # MPD is added

# model 3

cand <- list()

cand[[1]] <- update(LMM.PARS[[3]], . ~ . + AR)

cand[[2]] <- update(LMM.PARS[[3]], . ~ . + EPR)

cand[[3]] <- update(LMM.PARS[[3]], . ~ . + EPD)

cand[[4]] <- update(LMM.PARS[[3]], . ~ . + FPR)

cand[[5]] <- update(LMM.PARS[[3]], . ~ . + DR)

sapply(cand, AIC) %>% data.frame

sapply(cand, AIC) %>% which.min

sapply(LMM.PARS, AIC)

LMM.PARS[[4]] <- cand[[2]] # EPR is added

# model 4

cand <- list()

cand[[1]] <- update(LMM.PARS[[4]], . ~ . + AR)

cand[[2]] <- update(LMM.PARS[[4]], . ~ . + EPD)

cand[[3]] <- update(LMM.PARS[[4]], . ~ . + FPR)

cand[[4]] <- update(LMM.PARS[[4]], . ~ . + DR)

sapply(cand, AIC) %>% data.frame

sapply(cand, AIC) %>% which.min

sapply(LMM.PARS, AIC)

LMM.PARS[[5]] <- cand[[2]] # EPD is added

# model 5

cand <- list()

cand[[1]] <- update(LMM.PARS[[5]], . ~ . + AR)

cand[[2]] <- update(LMM.PARS[[5]], . ~ . + FPR)

cand[[3]] <- update(LMM.PARS[[5]], . ~ . + DR)

sapply(cand, AIC) %>% data.frame

sapply(cand, AIC) %>% which.min

sapply(LMM.PARS, AIC)

LMM.PARS[[6]] <- cand[[1]] # AR is added

# model 6

cand <- list()

cand[[1]] <- update(LMM.PARS[[6]], . ~ . + FPR)

cand[[2]] <- update(LMM.PARS[[6]], . ~ . + DR)

sapply(cand, AIC) %>% data.frame

sapply(cand, AIC) %>% which.min

sapply(LMM.PARS, AIC)

LMM.PARS[[7]] <- cand[[1]] # FPR is added

# model 7

cand <- list()

cand[[1]] <- update(LMM.PARS[[7]], . ~ . + DR)

sapply(cand, AIC) %>% data.frame

sapply(cand, AIC) %>% which.min

sapply(LMM.PARS, AIC) #AIC increased here

#Likelihood ratio test

anova(LMM.PARS[[1]],LMM.PARS[[2]],test="chi")

anova(LMM.PARS[[2]],LMM.PARS[[3]],test="chi")

anova(LMM.PARS[[3]],LMM.PARS[[4]],test="chi")

anova(LMM.PARS[[4]],LMM.PARS[[5]],test="chi")

anova(LMM.PARS[[5]],LMM.PARS[[6]],test="chi")

LMM.PARS[[5]]@call

summary(LMM.PARS[[5]])

r.squaredGLMM(LMM.PARS[[5]])

CI.ALL <- confint(LMM.PARS[[5]],method="Wald")

CI.ALL

check\_model(LMM.PARS[[5]])

vif(LMM.PARS[[5]])

### Modeling PF w/ MWS variables added to parsimonious UF model ###

# model 6

cand <- list()

cand[[1]] <- update(LMM.PARS[[5]], . ~ . + BiF.N)

cand[[2]] <- update(LMM.PARS[[5]], . ~ . + BiMI.N)

cand[[3]] <- update(LMM.PARS[[5]], . ~ . + BiPro)

cand[[4]] <- update(LMM.PARS[[5]], . ~ . + TriF.N)

cand[[5]] <- update(LMM.PARS[[5]], . ~ . + TriMI.N)

cand[[6]] <- update(LMM.PARS[[5]], . ~ . + TriPro)

sapply(cand, AIC) %>% data.frame

sapply(cand, AIC) %>% which.min

sapply(LMM.PARS, AIC)

LMM.PARS[[6]] <- cand[[3]] # BiPro is added

# model 7

cand <- list()

cand[[1]] <- update(LMM.PARS[[6]], . ~ . + BiF.N)

cand[[2]] <- update(LMM.PARS[[6]], . ~ . + BiMI.N)

cand[[3]] <- update(LMM.PARS[[6]], . ~ . + TriF.N)

cand[[4]] <- update(LMM.PARS[[6]], . ~ . + TriMI.N)

cand[[5]] <- update(LMM.PARS[[6]], . ~ . + TriPro)

sapply(cand, AIC) %>% data.frame

sapply(cand, AIC) %>% which.min

sapply(LMM.PARS, AIC)

LMM.PARS[[7]] <- cand[[2]] # BiMI.N is added

# model 8

cand <- list()

cand[[1]] <- update(LMM.PARS[[7]], . ~ . + BiF.N)

cand[[2]] <- update(LMM.PARS[[7]], . ~ . + TriF.N)

cand[[3]] <- update(LMM.PARS[[7]], . ~ . + TriMI.N)

cand[[4]] <- update(LMM.PARS[[7]], . ~ . + TriPro)

sapply(cand, AIC) %>% data.frame

sapply(cand, AIC) %>% which.min

sapply(LMM.PARS, AIC) #AIC increses here

#Likelihood ratio test

anova(LMM.PARS[[5]],LMM.PARS[[6]],test="chi")

anova(LMM.PARS[[6]],LMM.PARS[[7]],test="chi") #non-significant

LMM.PARS[[6]]@call

summary(LMM.PARS[[6]])

r.squaredGLMM(LMM.PARS[[6]])

CI.ALL <- confint(LMM.PARS[[6]],method="Wald")

CI.ALL

check\_model(LMM.PARS[[6]])

vif(LMM.PARS[[6]])

r.squaredGLMM(LMM.PARS[[1]])

r.squaredGLMM(LMM.PARS[[2]])

r.squaredGLMM(LMM.PARS[[3]])

r.squaredGLMM(LMM.PARS[[4]])

r.squaredGLMM(LMM.PARS[[5]])

r.squaredGLMM(LMM.PARS[[6]])