# Supplementary Information

## Procedure of selecting the picture stimuli for the picture naming task

A picture naming task is assumed to draw on two cognitive processes: (a) picture recognition (identifying the concept corresponding to the visual stimulus) and (b) name retrieval (retrieving the lemma in the target language corresponding to the concept identified). Therefore, to reduce the demands on picture recognition, care was taken to select an appropriate set of picture stimuli. First, target picture stimuli and names were pooled from Snodgrass and Vanderwart's (1980) list which has been used in various L1 and L2 studies (e.g., Leonard & Shea, 2017). Second, to select picture names familiar for the participants, the threshold of frequency level was determined as most frequent 3000 words in the JACET8000 wordlist (JACET, 2003). Third, two indices reported in Snodgrass and Vanderwart’s original study were considered to minimize variability in picture recognition difficulty: (a) name agreement (> 95%) and (b) image agreement (> 3.5 on a 5-point scale). In addition, we considered the indices of familiarity and visual complexity concerning Japanese culture. The final set of picture stimuli for the study included the 50 pictures below:

*Onion; Eye; Apple; Fish Pencil; Banana; Foot ;Bed; Pipe; Belt; Football Rabbit; Ruler; Fork; Book; Frog; Sandwich; Bowl; Shirt; Glass; Bus; Snake; Butterfly Glove; Sock; Button Guitar; Carrot; Hammer Star; Hat; Sun; Cat; Swing; Heart; Chain; Horse; Thumb; Door; Tree; Iron; Drum; Turtle; Key; Duck; Lemon Umbrella; Ear; Elephant Nose*

## Descriptive statistics of cognitive fluency and utterance fluency measures

Table S1. *Descriptive summary of cognitive fluency measures.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   |   |   |   | *Shapiro-Wilk test* |
| Cognitive fluency measures | *Mean* | *SD* | *SE* | *Statistics* | *p-value* |
| PVLT | 25.55 | 6.92 | 0.61 | 0.993 | 0.756 |
| Picture Naming RT | 1099.05 | 180.48 | 15.95 | 0.990 | 0.522 |
| Maze Word RT | 1164.43 | 202.93 | 17.94 | 0.969 | 0.005 |
| Maze Word Accuracy | 385.45 | 39.73 | 3.51 | 0.861 | < .001 |
| GJT Morphology RT | 4039.66 | 969.44 | 85.69 | 0.989 | 0.396 |
| GJT Syntax RT | 4291.60 | 955.92 | 84.49 | 0.990 | 0.461 |
| GJT Morphology Accuracy | 20.17 | 3.12 | 0.28 | 0.976 | 0.024 |
| GJT Syntax Accuracy | 31.28 | 4.13 | 0.37 | 0.947 | < .001 |
| Articulatory speed | 190.26 | 26.04 | 2.30 | 0.990 | 0.502 |
| *N.B.* RT measures are expressed in milliseconds. Articulatory speed refers to the mean number of morae per minute. |

Table S2*. Descriptive statistics and Shapiro-Wilk tests.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |   |   |   |   | *Shapiro-Wilk test* |
| UF measures | Task | *M* | *SD* | *SE* | *Statistics* | *p-value* |
| Articulation rate | Arg | 3.140 | 0.598 | 0.053 | 0.980 | 0.059 |
|  | PicN | 2.824 | 0.506 | 0.045 | 0.986 | 0.201 |
|  | RtoS | 2.664 | 0.459 | 0.041 | 0.990 | 0.526 |
|  | RwLtoS | 2.697 | 0.463 | 0.041 | 0.984 | 0.131 |
| Mid-clause pause ratio | Arg | 0.215 | 0.090 | 0.008 | 0.970 | 0.006 |
|  | PicN | 0.242 | 0.102 | 0.009 | 0.976 | 0.025 |
|  | RtoS | 0.255 | 0.104 | 0.009 | 0.904 | < .001 |
|   | RwLtoS | 0.259 | 0.093 | 0.008 | 0.982 | 0.080 |
| End-clause pause ratio | Arg | 0.059 | 0.021 | 0.002 | 0.968 | 0.004 |
|  | PicN | 0.086 | 0.023 | 0.002 | 0.991 | 0.603 |
|  | RtoS | 0.076 | 0.019 | 0.002 | 0.984 | 0.141 |
|   | RwLtoS | 0.076 | 0.019 | 0.002 | 0.968 | 0.004 |
| Filled pause ratio | Arg | 0.111 | 0.099 | 0.009 | 0.836 | < .001 |
|  | PicN | 0.097 | 0.094 | 0.008 | 0.843 | < .001 |
|  | RtoS | 0.116 | 0.108 | 0.010 | 0.803 | < .001 |
|   | RwLtoS | 0.118 | 0.095 | 0.008 | 0.876 | < .001 |
| Mid-clause pause duration | Arg | 1.120 | 0.522 | 0.046 | 0.790 | < .001 |
|  | PicN | 1.119 | 0.489 | 0.043 | 0.864 | < .001 |
|  | RtoS | 1.140 | 0.503 | 0.044 | 0.714 | < .001 |
|   | RwLtoS | 1.195 | 0.466 | 0.041 | 0.884 | < .001 |
| End-clause pause duration | Arg | 1.368 | 1.193 | 0.105 | 0.608 | < .001 |
|  | PicN | 1.302 | 0.732 | 0.065 | 0.855 | < .001 |
|  | RtoS | 1.590 | 1.334 | 0.118 | 0.605 | < .001 |
|   | RwLtoS | 1.485 | 1.178 | 0.104 | 0.531 | < .001 |
| Self-repetition ratio | Arg | 0.076 | 0.070 | 0.006 | 0.858 | < .001 |
|  | PicN | 0.102 | 0.072 | 0.006 | 0.918 | < .001 |
|  | RtoS | 0.090 | 0.079 | 0.007 | 0.838 | < .001 |
|   | RwLtoS | 0.094 | 0.072 | 0.006 | 0.853 | < .001 |
| Self-correction ratio | Arg | 0.021 | 0.016 | 0.001 | 0.923 | < .001 |
|  | PicN | 0.025 | 0.017 | 0.002 | 0.943 | < .001 |
|  | RtoS | 0.025 | 0.019 | 0.002 | 0.827 | < .001 |
|   | RwLtoS | 0.024 | 0.016 | 0.001 | 0.954 | < .001 |
| False start ratio | Arg | 0.008 | 0.011 | 0.001 | 0.735 | < .001 |
|  | PicN | 0.007 | 0.008 | 0.001 | 0.795 | < .001 |
|  | RtoS | 0.012 | 0.012 | 0.001 | 0.853 | < .001 |
|   | RwLtoS | 0.011 | 0.010 | 0.001 | 0.890 | < .001 |

*Note*. Arg = Argumentative task; PicN = Picture narrative task; RtoS = Reading-to-speaking task; RwLtoS = Reading-while-Listening to speaking task.

Table S3. *A correlational matrix of utterance fluency measures and cognitive fluency measures across tasks.*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Task | AR | SR | MLR | MCPR | ECPR | FPR | MCPD | ECPD | SRR | SCR | FSR |
| PVLT | Arg | 0.395\*\*\* | 0.425\*\*\* | 0.434\*\*\* | -0.430\*\*\* | -0.318\*\*\* | -0.302\*\*\* | -0.236\*\* | -0.324\*\*\* | -0.195\* | -0.073 | -0.135 |
|  | PicN | 0.400\*\*\* | 0.415\*\*\* | 0.459\*\*\* | -0.442\*\*\* | -0.323\*\*\* | -0.324\*\*\* | -0.267\*\* | -0.271\*\* | -0.215\* | -0.117 | -0.107 |
|  | RtoS | 0.421\*\*\* | 0.481\*\*\* | 0.506\*\*\* | -0.508\*\*\* | -0.326\*\*\* | -0.370\*\*\* | -0.276\*\* | -0.363\*\*\* | -0.244\*\* | -0.251\*\* | -0.017 |
|   | RwLtoS | 0.363\*\*\* | 0.398\*\*\* | 0.406\*\*\* | -0.393\*\*\* | -0.238\*\* | -0.398\*\*\* | -0.250\*\* | -0.219\* | -0.167 | -0.169 | -0.070 |
| Picture Naming RT | Arg | -0.398\*\*\* | -0.455\*\*\* | -0.377\*\*\* | 0.404\*\*\* | 0.113 | 0.264\*\* | 0.302\*\*\* | 0.323\*\*\* | 0.273\*\* | 0.115 | 0.055 |
|  | PicN | -0.333\*\*\* | -0.401\*\*\* | -0.329\*\*\* | 0.333\*\*\* | 0.121 | 0.166 | 0.343\*\*\* | 0.323\*\*\* | 0.234\*\* | -0.001 | 0.182\* |
|  | RtoS | -0.346\*\*\* | -0.448\*\*\* | -0.378\*\*\* | 0.392\*\*\* | 0.078 | 0.161 | 0.368\*\*\* | 0.401\*\*\* | 0.150 | -0.017 | 0.035 |
|   | RwLtoS | -0.332\*\*\* | -0.450\*\*\* | -0.371\*\*\* | 0.365\*\*\* | 0.243\*\* | 0.217\* | 0.359\*\*\* | 0.308\*\*\* | 0.238\*\* | -0.053 | 0.087 |
| Maze Word RT | Arg | -0.356\*\*\* | -0.400\*\*\* | -0.379\*\*\* | 0.392\*\*\* | 0.141 | 0.287\*\* | 0.253\*\* | 0.303\*\*\* | 0.168 | -0.001 | 0.044 |
|  | PicN | -0.424\*\*\* | -0.464\*\*\* | -0.451\*\*\* | 0.446\*\*\* | 0.186\* | 0.275\*\* | 0.346\*\*\* | 0.325\*\*\* | 0.226\* | 0.063 | 0.254\*\* |
|  | RtoS | -0.418\*\*\* | -0.501\*\*\* | -0.469\*\*\* | 0.499\*\*\* | 0.168 | 0.281\*\* | 0.361\*\*\* | 0.393\*\*\* | 0.248\*\* | 0.101 | 0.058 |
|   | RwLtoS | -0.327\*\*\* | -0.409\*\*\* | -0.428\*\*\* | 0.427\*\*\* | 0.244\*\* | 0.266\*\* | 0.242\*\* | 0.236\*\* | 0.131 | 0.022 | 0.115 |
| Maze Word Acc | Arg | 0.168 | 0.180\* | 0.197\* | -0.198\* | -0.123 | -0.204\* | -0.115 | -0.156 | -0.087 | 0.056 | -0.183\* |
|  | PicN | 0.187\* | 0.204\* | 0.247\*\* | -0.263\*\* | -0.052 | -0.261\*\* | -0.168 | -0.105 | -0.169 | -0.079 | -0.194\* |
|  | RtoS | 0.208\* | 0.235\*\* | 0.228\*\* | -0.260\*\* | -0.089 | -0.242\*\* | -0.158 | -0.205\* | -0.167 | -0.193\* | -0.047 |
|   | RwLtoS | 0.177\* | 0.175\* | 0.165 | -0.174\* | 0.023 | -0.277\*\* | -0.184\* | -0.062 | -0.197\* | -0.173 | -0.064 |
| GJT Morph RT | Arg | -0.327\*\*\* | -0.356\*\*\* | -0.370\*\*\* | 0.361\*\*\* | 0.240\*\* | 0.160 | 0.256\*\* | 0.103 | 0.190\* | 0.133 | 0.175\* |
|  | PicN | -0.402\*\*\* | -0.403\*\*\* | -0.392\*\*\* | 0.346\*\*\* | 0.323\*\*\* | 0.097 | 0.229\*\* | 0.302\*\*\* | 0.193\* | 0.101 | 0.074 |
|  | RtoS | -0.327\*\*\* | -0.392\*\*\* | -0.402\*\*\* | 0.357\*\*\* | 0.333\*\*\* | 0.143 | 0.272\*\* | 0.322\*\*\* | 0.146 | 0.111 | -0.041 |
|   | RwLtoS | -0.275\*\* | -0.402\*\*\* | -0.378\*\*\* | 0.351\*\*\* | 0.310\*\*\* | 0.131 | 0.302\*\*\* | 0.302\*\*\* | 0.054 | 0.158 | 0.023 |
| GJT Syn RT | Arg | -0.351\*\*\* | -0.394\*\*\* | -0.415\*\*\* | 0.407\*\*\* | 0.255\*\* | 0.222\* | 0.288\*\* | 0.130 | 0.208\* | 0.124 | 0.133 |
|  | PicN | -0.428\*\*\* | -0.444\*\*\* | -0.440\*\*\* | 0.386\*\*\* | 0.389\*\*\* | 0.169 | 0.273\*\* | 0.325\*\*\* | 0.249\*\* | 0.121 | 0.049 |
|  | RtoS | -0.355\*\*\* | -0.432\*\*\* | -0.448\*\*\* | 0.403\*\*\* | 0.366\*\*\* | 0.195\* | 0.317\*\*\* | 0.340\*\*\* | 0.193\* | 0.108 | 0.001 |
|   | RwLtoS | -0.271\*\* | -0.427\*\*\* | -0.423\*\*\* | 0.391\*\*\* | 0.343\*\*\* | 0.171 | 0.299\*\*\* | 0.322\*\*\* | 0.070 | 0.116 | 0.010 |
| GJT MorphAcc | Arg | 0.170 | 0.149 | 0.142 | -0.142 | -0.203\* | -0.095 | -0.051 | -0.168 | -0.076 | -0.068 | 0.017 |
|  | PicN | 0.161 | 0.143 | 0.127 | -0.122 | -0.142 | -0.070 | -0.138 | -0.093 | 0.038 | 0.021 | -0.183\* |
|  | RtoS | 0.208\* | 0.189\* | 0.181\* | -0.199\* | -0.080 | -0.118 | -0.114 | -0.177\* | -0.074 | -0.004 | -0.078 |
|   | RwLtoS | 0.187\* | 0.125 | 0.159 | -0.154 | -0.090 | -0.106 | -0.053 | 0.012 | -0.084 | -0.068 | -0.017 |
| GJT Syn Acc | Arg | 0.321\*\*\* | 0.345\*\*\* | 0.399\*\*\* | -0.404\*\*\* | -0.261\*\* | -0.251\*\* | -0.138 | -0.284\*\* | -0.168 | -0.065 | -0.121 |
|  | PicN | 0.302\*\*\* | 0.355\*\*\* | 0.396\*\*\* | -0.398\*\*\* | -0.174 | -0.354\*\*\* | -0.300\*\*\* | -0.211\* | -0.199\* | -0.156 | -0.228\*\* |
|  | RtoS | 0.313\*\*\* | 0.393\*\*\* | 0.374\*\*\* | -0.414\*\*\* | -0.130 | -0.319\*\*\* | -0.293\*\*\* | -0.300\*\*\* | -0.297\*\*\* | -0.208\* | -0.070 |
|   | RwLtoS | 0.270\*\* | 0.278\*\* | 0.332\*\*\* | -0.314\*\*\* | -0.190\* | -0.315\*\*\* | -0.175\* | -0.140 | -0.208\* | -0.043 | -0.084 |
| Articulatory speed | Arg | 0.546\*\*\* | 0.470\*\*\* | 0.455\*\*\* | -0.459\*\*\* | -0.190\* | -0.150 | -0.281\*\* | -0.233\*\* | -0.158 | -0.125 | -0.075 |
|  | PicN | 0.601\*\*\* | 0.528\*\*\* | 0.501\*\*\* | -0.494\*\*\* | -0.249\*\* | -0.136 | -0.377\*\*\* | -0.353\*\*\* | -0.218\* | -0.061 | -0.174\* |
|  | RtoS | 0.557\*\*\* | 0.576\*\*\* | 0.522\*\*\* | -0.533\*\*\* | -0.153 | -0.214\* | -0.449\*\*\* | -0.388\*\*\* | -0.235\*\* | -0.131 | -0.051 |
|   | RwLtoS | 0.533\*\*\* | 0.534\*\*\* | 0.527\*\*\* | -0.537\*\*\* | -0.261\*\* | -0.150 | -0.328\*\*\* | -0.322\*\*\* | -0.071 | 0.062 | -0.229\*\* |

*Note*. \* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001; PVLT = Productive Vocabulary Levels Test; Maze Word Acc = Maze Word accuracy; GJT Morph RT = GJT Morphology RT; GJT Syn RT = GJT Syntax RT; GJT Morph Acc = GJT Morphology Accuracy; GJT Syn Acc = GJT Syntax Accuracy

## The parameters of the final CFA model of cognitive fluency

Table S4. *Summary of the standardized regression coefficients and their 95% confidence intervals of the finalized CFA model of cognitive fluency*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|   |   |   |   |   | 95%CI |
| Latent variable | Direction | Observed variable | *β* | *p* | Lower | Upper |
| ***Covariance between latent variables*** |  |  |  |  |
| Linguistic resource | **–** | Processing speed | 0.676 | < .001 | 0.515 | 0.838 |
| ***Measurement model*** |  |  |  |  |  |  |
| Linguistic resource | → | PVLT | 0.867 | < .001 | 0.800 | 0.933 |
|  | → | Maze Word Accuracy | 0.689 | < .001 | 0.552 | 0.826 |
|  | → | GJT Morph. Accuracy | 0.455 | < .001 | 0.264 | 0.647 |
|  | → | GJT Syn. Accuracy | 0.722 | < .001 | 0.610 | 0.834 |
| Processing speed | → | Pic. Naming RT | 0.424 | < .001 | 0.269 | 0.580 |
|  | → | Maze Word RT | 0.862 | < .001 | 0.762 | 0.963 |
|  | → | GJT Morph. RT | 0.616 | < .001 | 0.472 | 0.761 |
|  | → | GJT Syn. RT | 0.590 | < .001 | 0.441 | 0.739 |
|   | → | Articulatory speed | 0.589 | < .001 | 0.464 | 0.714 |

## Procedures of model modification of CFA models of utterance fluency (with the first three UF models)

The intercorrelation pooled by tasks was tested through parametric correlational analyses (i.e., Pearson product-moment correlation). This is because an SEM analysis is based on the correlation-matrix based on the parametric correlation coefficients. To inspect the overall intercollinearity among the UF measures, the dataset was pooled over tasks. The correlation coefficients and their heatmap are presented below respectively as Table S5 and Figure S1. To avoid strong collinearity among observed variables, strong correlations, particularly across latent variables (e.g., speed and breakdown fluency measures), were excluded. According to the correlation matrix and the heatmap visualization, speech rate strongly correlated with mid-clause pause ratio (breakdown fluency; *r* = .845) and articulation rate (speed fluency; *r* = .859). Although mean length of run also indicated the strong correlations with articulation rate and mid-clause pause ratio, mean length of run showed a relatively weaker correlation with mid-clause pause ratio (*r* = .731). Considering the fact that both speech rate and mean length of run were the observed variables loaded onto the latent variable of speed fluency, mean length of run would result in relatively weak collinearity, compared to speech rate. In addition, within the observed variables of breakdown fluency, mid-clause pause duration and end-clause pause duration were strongly correlated with each other (*r* = .735). Although mid- and end-clause pauses are theoretically supposed to represent different underlying processing (De Jong, 2016; Tavakoli, 2011), the role of pause location in duration may not be statistically distinctive in factor analyses. Therefore, we decided to exclude speech rate from the measurement model of speed fluency and to replace mid-clause pause duration and end-clause pause duration with the mean pause duration measure which was calculated as the mean duration of pauses including both end- and mid-clause pauses. The revised correlation matrix is presented in Table S6.



*Figure S1.* The heatmap visualization of correlation coefficients between utterance fluency measures.

*Note*. \* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001; Each cell refers to each data point of the correlation matrix, and the values in the cells are the correlation coefficients of the data points. The thickness of colour of the cells indicates the strengths of correlation coefficients, meaning that the thicker purple the cell is, the stronger correlation coefficient it shows.

Table S5. *A correlational matrix of the utterance fluency measures pooled across four tasks.*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2. SR | 3. MLR | 4. MCPR | 5. ECPR | 6. FPR | 7. MCPD | 8. ECPD | 9. SRR | 10. SCR | 11. FSR |
| 1. AR | 0.859\*\*\* | 0.714\*\*\* | 0.760\*\*\* | 0.400\*\*\* | 0.364\*\*\* | 0.480\*\*\* | 0.375\*\*\* | 0.533\*\*\* | 0.375\*\*\* | 0.370\*\*\* |
| 2. SR | — | 0.839\*\*\* | 0.845\*\*\* | 0.485\*\*\* | 0.499\*\*\* | 0.703\*\*\* | 0.555\*\*\* | 0.488\*\*\* | 0.297\*\*\* | 0.281\*\*\* |
| 3. MLR |  | — | 0.731\*\*\* | 0.473\*\*\* | 0.371\*\*\* | 0.386\*\*\* | 0.306\*\*\* | 0.399\*\*\* | 0.305\*\*\* | 0.228\*\*\* |
| 4. MCPR |  |  | — | 0.362\*\*\* | 0.557\*\*\* | 0.604\*\*\* | 0.465\*\*\* | 0.586\*\*\* | 0.409\*\*\* | 0.323\*\*\* |
| 5. ECPR |  |  |  | — | 0.204\*\*\* | 0.226\*\*\* | 0.277\*\*\* | 0.186\*\*\* | 0.209\*\*\* | 0.079 |
| 6. FPR |  |  |  |  | — | 0.545\*\*\* | 0.430\*\*\* | 0.497\*\*\* | 0.295\*\*\* | 0.236\*\*\* |
| 7. MCPD |  |  |  |  |  | — | 0.735\*\*\* | 0.344\*\*\* | 0.120\*\* | 0.151\*\*\* |
| 8. ECPD |  |  |  |  |  |  | — | 0.289\*\*\* | 0.098\* | 0.025 |
| 9. SRR |  |  |  |  |  |  |  | — | 0.499\*\*\* | 0.328\*\*\* |
| 10. SCR |  |  |  |  |  |  |  |  | — | 0.159\*\*\* |
| 11. FSR |   |   |   |   |   |   |   |   |   | — |

*Note*. \* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001; AR = Articulation rate; SR = Speech rate; MLR = Mean length of run; MCPR = Mid-clause pause ratio; ECPR = End-clause pause ratio; FPR = Filled pause ratio; MCPD = Mid-clause pause duration; ECPD = End-clause pause duration; SRR = Self-repetition ratio; SCR = Self-correction ratio; FSR = False start ratio.

Table S6. *A revised correlational matrix of the utterance fluency measures pooled across four tasks.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 2. MLR | 3. MCPR | 4. ECPR | 5. FPR | 6. MPD | 7. SRR | 8. SCR | 9. FSR |
| 1. AR | 0.714\*\*\* | 0.760\*\*\* | 0.400\*\*\* | 0.364\*\*\* | 0.469\*\*\* | 0.533\*\*\* | 0.375\*\*\* | 0.370\*\*\* |
| 2. MLR | — | 0.731\*\*\* | 0.473\*\*\* | 0.371\*\*\* | 0.386\*\*\* | 0.399\*\*\* | 0.305\*\*\* | 0.228\*\*\* |
| 3. MCPR |  | — | 0.362\*\*\* | 0.557\*\*\* | 0.577\*\*\* | 0.586\*\*\* | 0.409\*\*\* | 0.323\*\*\* |
| 4. ECPR |  |  | — | 0.204\*\*\* | 0.285\*\*\* | 0.186\*\*\* | 0.209\*\*\* | 0.079 |
| 5. FPR |  |  |  | — | 0.532\*\*\* | 0.497\*\*\* | 0.295\*\*\* | 0.236\*\*\* |
| 6. MPD |  |  |  |  | — | 0.342\*\*\* | 0.113\* | 0.106\* |
| 7. SRR |  |  |  |  |  | — | 0.499\*\*\* | 0.328\*\*\* |
| 8. SCR |  |  |  |  |  |  | — | 0.159\*\*\* |
| 9. FSR |   |   |   |   |   |   |   | — |

*Note*. \* *p* < .05, \*\* *p* < .01, \*\*\* *p* < .001; AR = Articulation rate; MLR = Mean length of run; MCPR = Mid-clause pause ratio; ECPR = End-clause pause ratio; FPR = Filled pause ratio; MPD = Mean pause duration; SRR = Self-repetition ratio; SCR = Self-correction ratio; FSR = False start ratio.

In addition to reducing the potential collinearity among observed variables, the modification indices were also calculated to explore some residuals that can be replaced with residual covariances to improve the model fit. However, the modification indices only statistically suggest the additional paths that can improve the model fit; the suggested paths were thus accepted only if the residual covariances can be theoretically explained (Raykov & Marcoulides, 2006; Tabachnick & Fidell, 1996). Eventually, the following three residual covariances were adopted.

First, the residual covariance between mean pause duration and filled pause ratio was considered justifiable because when speakers produced relatively longer pauses, they were likely to utilize filled pauses to provide the impression of continuation of speech (Clark & Fox Tree, 2002). This is also supported by a moderate correlation between them in the current dataset (*r* = .532, see Table S6). Accordingly, some use of filled pauses may be derived from speakers’ speaking strategies used for making their speech sound more fluent (Clark & Fox Tree, 2002). In other words, the residual covariance between mean pause duration and filled pause ratio may come from some common idiosyncratic factors other than the construct of breakdown fluency.

Second, the residual covariance between mid-clause pause ratio and self-correction ratio was accepted. From the perspective of speech production mechanisms, mid-clause pauses represent the disruptions in speech processing due to the lack of linguistic resources (De Jong, 2016; Götz, 2013; Tavakoli, 2011), whereas self-repairs are supposed to indicate overt monitoring processes (Kormos, 2000, 2006). Accordingly, the residual covariance between these two measures can be theoretically explained; when speakers produce breakdowns in the middle of clauses due to the lack of particular linguistic knowledge, they are usually required to maintain their fluency by modifying their utterances. This possible pattern of self-corrections triggered by mid-clause pauses was also supported by the moderate correlation in the current study (*r* = .409, see Table S6). Since these two measures belong to different constructs (breakdown and repair fluency, respectively), this shared residual was illustrated as a residual covariance.

Third, the residual covariance between end-clause pause ratio and false start ratio was adopted, because when speakers produce false starts, they are supposed to be engaged with conceptualization processes (Williams & Korko, 2019), which end-clause pauses are also supposed to reflect (De Jong, 2016; Tavakoli, 2011). More specifically, speakers correct their utterances at the beginning of the utterance (i.e., false start) for the sake of content information appropriacy or correctness. From a theoretical perspective, such content information is specified by conceptualization processes. Therefore, it seems plausible to argue that the causes of both end-clause pauses and false starts are associated with high demands on conceptualization processes (breakdowns vs. overt monitoring for content planning). As with the second residual covariance, this shared residual across constructs was included in the CFA models as a residual covariance.

## The parameters of the final CFA model of utterance fluency

Table S7. *Summary of the standardized regression coefficients and their 95% confidence intervals of the three-factor CFA model of cognitive fluency.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|   |   |   |   |   |   | 95%CI |
| Latent variable | Direction | Observed variable | Task | *β* | *p* | Lower | Upper |
| ***Covariance between latent variables*** |  |  |  |  |  |  |
| Speed fluency | vs. | Breakdown fluency | Arg | 0.929 | < .001 | 0.845 | 1.014 |
|  |  |  | PicN | 0.960 | < .001 | 0.910 | 1.011 |
|  |  |  | RtoS | 0.948 | < .001 | 0.868 | 1.028 |
|  |  |  | RwLtoS | 0.945 | < .001 | 0.879 | 1.011 |
| Speed fluency | vs. | Repair fluency | Arg | 0.732 | < .001 | 0.607 | 0.857 |
|  |  |  | PicN | 0.704 | < .001 | 0.603 | 0.804 |
|  |  |  | RtoS | 0.627 | < .001 | 0.489 | 0.766 |
|  |  |  | RwLtoS | 0.600 | < .001 | 0.445 | 0.755 |
| Breakdown fluency | vs. | Repair fluency | Arg | 0.819 | < .001 | 0.688 | 0.951 |
|  |  |  | PicN | 0.732 | < .001 | 0.616 | 0.848 |
|  |  |  | RtoS | 0.762 | < .001 | 0.610 | 0.914 |
|  |  |  | RwLtoS | 0.621 | < .001 | 0.458 | 0.785 |
| ***Measurement model*** |  |  |  |  |  |  |  |
| Speed fluency | → | Articulation rate | Arg | 0.863 | < .001 | 0.770 | 0.957 |
|  |  |  | PicN | 0.865 | < .001 | 0.803 | 0.927 |
|  |  |  | RtoS | 0.822 | < .001 | 0.737 | 0.908 |
|  |  |  | RwLtoS | 0.803 | < .001 | 0.707 | 0.899 |
|  | → | Mean length or run | Arg | 0.768 | < .001 | 0.693 | 0.844 |
|  |  |  | PicN | 0.926 | < .001 | 0.903 | 0.949 |
|  |  |  | RtoS | 0.926 | < .001 | 0.889 | 0.963 |
|  |  |  | RwLtoS | 0.911 | < .001 | 0.878 | 0.945 |
| Breakdown fluency | → | Mid-clause pause ratio | Arg | 0.979 | < .001 | 0.931 | 1.027 |
|  |  |  | PicN | 0.960 | < .001 | 0.929 | 0.991 |
|  |  |  | RtoS | 0.950 | < .001 | 0.876 | 1.025 |
|  |  |  | RwLtoS | 0.982 | < .001 | 0.934 | 1.030 |
|  | → | End-clause pause ratio | Arg | 0.471 | < .001 | 0.315 | 0.628 |
|  |  |  | PicN | 0.347 | < .001 | 0.203 | 0.490 |
|  |  |  | RtoS | 0.399 | < .001 | 0.232 | 0.566 |
|  |  |  | RwLtoS | 0.434 | < .001 | 0.276 | 0.591 |
|  | → | Mean pause duration | Arg | 0.499 | < .001 | 0.389 | 0.610 |
|  |  |  | PicN | 0.658 | < .001 | 0.559 | 0.756 |
|  |  |  | RtoS | 0.621 | < .001 | 0.431 | 0.811 |
|  |  |  | RwLtoS | 0.543 | < .001 | 0.335 | 0.752 |
|  | → | Filled pause ratio | Arg | 0.601 | < .001 | 0.476 | 0.725 |
|  |  |  | PicN | 0.564 | < .001 | 0.445 | 0.684 |
|  |  |  | RtoS | 0.570 | < .001 | 0.362 | 0.779 |
|  |  |  | RwLtoS | 0.548 | < .001 | 0.399 | 0.696 |
| Repair fluency | → | False start ratio | Arg | 0.445 | < .001 | 0.244 | 0.645 |
|  |  |  | PicN | 0.473 | < .001 | 0.299 | 0.648 |
|  |  |  | RtoS | 0.281 | 0.011 | 0.066 | 0.497 |
|  |  |  | RwLtoS | 0.415 | < .001 | 0.208 | 0.623 |
|  | → | Self-repetition ratio | Arg | 0.833 | < .001 | 0.729 | 0.936 |
|  |  |  | PicN | 0.825 | < .001 | 0.731 | 0.920 |
|  |  |  | RtoS | 0.877 | < .001 | 0.753 | 1.001 |
|  |  |  | RwLtoS | 0.812 | < .001 | 0.622 | 1.001 |
|  | → | Self-correction ratio | Arg | 0.580 | < .001 | 0.423 | 0.737 |
|  |  |  | PicN | 0.632 | < .001 | 0.506 | 0.757 |
|  |  |  | RtoS | 0.591 | < .001 | 0.359 | 0.822 |
|   |   |   | RwLtoS | 0.475 | < .001 | 0.314 | 0.635 |

## The parameters of the final SEM model of the cognitive-utterance fluency link

Table S8. *Summary of the standardized regression coefficients and their 95% confidence intervals of the structural model of cognitive fluency and utterance fluency.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|   |   |   |   |   |   | 95%CI |
| Latent variable | Direction | Latent variable | Task | *β* | *p* | Lower | Upper |
| ***Regression model*** |  |  |  |  |  |  |  |
| Linguistic resource | → | Speed fluency | Arg | 0.168 | 0.061 | -0.008 | 0.344 |
|  |  |  | PicN | 0.161 | 0.104 | -0.033 | 0.354 |
|  |  |  | RtoS | 0.234 | 0.038 | 0.013 | 0.455 |
|  |   |   | RwLtoS | 0.276 | 0.004 | 0.086 | 0.465 |
|  | → | Breakdown fluency | Arg | 0.345 | 0.001 | 0.139 | 0.550 |
|  |  |  | PicN | 0.240 | 0.025 | 0.030 | 0.451 |
|  |  |  | RtoS | 0.317 | 0.014 | 0.064 | 0.570 |
|  |   |   | RwLtoS | 0.221 | 0.061 | -0.010 | 0.452 |
|  | → | Repair fluency | Arg | 0.225 | 0.150 | -0.081 | 0.531 |
|  |  |  | PicN | 0.330 | 0.049 | 0.002 | 0.659 |
|  |  |  | RtoS | 0.375 | 0.019 | 0.062 | 0.689 |
|   |   |   | RwLtoS | 0.360 | 0.033 | 0.029 | 0.692 |
| Processing speed | → | Speed fluency | Arg | 0.533 | < .001 | 0.373 | 0.693 |
|  |  |  | PicN | 0.609 | < .001 | 0.434 | 0.784 |
|  |  |  | RtoS | 0.566 | < .001 | 0.371 | 0.761 |
|  |   |   | RwLtoS | 0.431 | < .001 | 0.244 | 0.617 |
|  | → | Breakdown fluency | Arg | 0.376 | < .001 | 0.191 | 0.561 |
|  |  |  | PicN | 0.501 | < .001 | 0.314 | 0.689 |
|  |  |  | RtoS | 0.411 | 0.003 | 0.144 | 0.679 |
|  |   |   | RwLtoS | 0.480 | < .001 | 0.251 | 0.710 |
|  | → | Repair fluency | Arg | 0.136 | 0.349 | -0.149 | 0.420 |
|  |  |  | PicN | 0.129 | 0.351 | -0.142 | 0.400 |
|  |  |  | RtoS | 0.094 | 0.452 | -0.152 | 0.341 |
|  |  |  | RwLtoS | -0.020 | 0.906 | -0.351 | 0.311 |

Table S9. *Summary of the standardized regression coefficients and their 95% confidence intervals of the measurement model of cognitive fluency in the final SEM model.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|   |   |   |   |   |   | 95%CI |
| Latent variable | Direction | Observed variable | Task | *β* | *p* | Lower | Upper |
| ***Measurement model of cognitive fluency*** |  |  |  |  |  |
| Linguistic resource | → | PVLT | Arg | 0.850 | < .001 | 0.788 | 0.912 |
|  |  |  | PicN | 0.845 | < .001 | 0.783 | 0.908 |
|  |  |  | RtoS | 0.870 | < .001 | 0.808 | 0.933 |
|  |   |   | RwLtoS | 0.879 | < .001 | 0.817 | 0.941 |
|  | → | Maze Word Accuracy | Arg | 0.680 | < .001 | 0.575 | 0.785 |
|  |  |  | PicN | 0.691 | < .001 | 0.573 | 0.810 |
|  |  |  | RtoS | 0.677 | < .001 | 0.549 | 0.805 |
|  |   |   | RwLtoS | 0.675 | < .001 | 0.552 | 0.799 |
|  | → | GJT Morph. Accuracy | Arg | 0.455 | < .001 | 0.305 | 0.605 |
|  |  |  | PicN | 0.439 | < .001 | 0.279 | 0.600 |
|  |  |  | RtoS | 0.441 | < .001 | 0.261 | 0.621 |
|  |   |   | RwLtoS | 0.442 | < .001 | 0.267 | 0.616 |
|  | → | GJT Syn. Accuracy | Arg | 0.746 | < .001 | 0.653 | 0.839 |
|  |  |  | PicN | 0.742 | < .001 | 0.648 | 0.836 |
|  |  |  | RtoS | 0.722 | < .001 | 0.617 | 0.828 |
|   |   |   | RwLtoS | 0.710 | < .001 | 0.605 | 0.814 |
| Processing speed | → | Pic. Naming RT | Arg | 0.450 | < .001 | 0.303 | 0.596 |
|  |  |  | PicN | 0.436 | < .001 | 0.298 | 0.573 |
|  |  |  | RtoS | 0.453 | < .001 | 0.295 | 0.611 |
|  |   |   | RwLtoS | 0.439 | < .001 | 0.289 | 0.589 |
|  | → | Maze Word RT | Arg | 0.821 | < .001 | 0.738 | 0.905 |
|  |  |  | PicN | 0.794 | < .001 | 0.702 | 0.887 |
|  |  |  | RtoS | 0.813 | < .001 | 0.716 | 0.909 |
|  |   |   | RwLtoS | 0.815 | < .001 | 0.719 | 0.912 |
|  | → | GJT Morph. RT | Arg | 0.617 | < .001 | 0.479 | 0.755 |
|  |  |  | PicN | 0.622 | < .001 | 0.482 | 0.762 |
|  |  |  | RtoS | 0.614 | < .001 | 0.480 | 0.748 |
|  |   |   | RwLtoS | 0.626 | < .001 | 0.498 | 0.754 |
|  | → | GJT Syn. RT | Arg | 0.604 | < .001 | 0.467 | 0.741 |
|  |  |  | PicN | 0.620 | < .001 | 0.487 | 0.754 |
|  |  |  | RtoS | 0.607 | < .001 | 0.473 | 0.741 |
|  |   |   | RwLtoS | 0.612 | < .001 | 0.489 | 0.736 |
|  | → | Articulatory speed | Arg | 0.635 | < .001 | 0.519 | 0.750 |
|  |  |  | PicN | 0.663 | < .001 | 0.551 | 0.774 |
|  |  |  | RtoS | 0.658 | < .001 | 0.543 | 0.773 |
|  |  |  | RwLtoS | 0.659 | < .001 | 0.544 | 0.775 |

Table S10. *Summary of the standardized regression coefficients between the latent variables of cognitive fluency and their 95% confidence intervals in the final SEM model.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|   |   |   |   |   |   | 95%CI |
| Latent variable | Direction | Latent variable | Task | *β* | *p* | Lower | Upper |
| ***Covariance between latent variables*** |  |  |  |  |  |
| Linguistic resource | vs. | Processing speed | Arg | 0.667 | < .001 | 0.515 | 0.819 |
|  |  |  | PicN | 0.664 | < .001 | 0.516 | 0.812 |
|  |  |  | RtoS | 0.671 | < .001 | 0.531 | 0.811 |
|  |   |   | RwLtoS | 0.676 | < .001 | 0.534 | 0.817 |

Table S11. *Summary of the standardized regression coefficients and their 95% confidence intervals of the measurement model of utterance fluency in the final SEM model.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|   |   |   |   |   |   | 95%CI |
| Latent variable | Direction | Observed variable | Task | *β* | *p* | Lower | Upper |
| ***Measurement model of utterance fluency*** |  |  |  |  |  |
| Speed fluency | → | Articulation rate | Arg | 0.905 | < .001 | 0.836 | 0.974 |
|  |  |  | PicN | 0.892 | < .001 | 0.838 | 0.947 |
|  |  |  | RtoS | 0.876 | < .001 | 0.805 | 0.948 |
|  |   |   | RwLtoS | 0.879 | < .001 | 0.803 | 0.955 |
|  | → | Mean length or run | Arg | 0.721 | < .001 | 0.632 | 0.810 |
|  |  |  | PicN | 0.882 | < .001 | 0.834 | 0.930 |
|  |  |  | RtoS | 0.831 | < .001 | 0.773 | 0.889 |
|   |   |   | RwLtoS | 0.800 | < .001 | 0.729 | 0.872 |
| Breakdown fluency | → | Mid-clause pause ratio | Arg | 0.958 | < .001 | 0.911 | 1.005 |
|  |  |  | PicN | 0.963 | < .001 | 0.922 | 1.004 |
|  |  |  | RtoS | 0.919 | < .001 | 0.832 | 1.005 |
|  |   |   | RwLtoS | 0.933 | < .001 | 0.877 | 0.990 |
|  | → | End-clause pause ratio | Arg | 0.515 | < .001 | 0.367 | 0.663 |
|  |  |  | PicN | 0.455 | < .001 | 0.310 | 0.601 |
|  |  |  | RtoS | 0.373 | < .001 | 0.205 | 0.540 |
|  |   |   | RwLtoS | 0.449 | < .001 | 0.233 | 0.664 |
|  | → | Mean pause duration | Arg | 0.528 | < .001 | 0.379 | 0.676 |
|  |  |  | PicN | 0.690 | < .001 | 0.595 | 0.786 |
|  |  |  | RtoS | 0.681 | < .001 | 0.499 | 0.862 |
|  |   |   | RwLtoS | 0.617 | < .001 | 0.407 | 0.827 |
|  | → | Filled pause ratio | Arg | 0.628 | < .001 | 0.510 | 0.746 |
|  |  |  | PicN | 0.545 | < .001 | 0.427 | 0.663 |
|  |  |  | RtoS | 0.598 | < .001 | 0.383 | 0.813 |
|   |   |   | RwLtoS | 0.556 | < .001 | 0.404 | 0.708 |
| Repair fluency | → | False starts ratio | Arg | 0.450 | < .001 | 0.281 | 0.619 |
|  |  |  | PicN | 0.459 | < .001 | 0.304 | 0.614 |
|  |  |  | RtoS | 0.289 | 0.014 | 0.059 | 0.518 |
|  |   |   | RwLtoS | 0.427 | < .001 | 0.235 | 0.620 |
|  | → | Self-repetition ratio | Arg | 0.827 | < .001 | 0.734 | 0.919 |
|  |  |  | PicN | 0.837 | < .001 | 0.756 | 0.917 |
|  |  |  | RtoS | 0.860 | < .001 | 0.747 | 0.973 |
|  |   |   | RwLtoS | 0.787 | < .001 | 0.648 | 0.925 |
|  | → | Self-correction ratio | Arg | 0.587 | < .001 | 0.453 | 0.721 |
|  |  |  | PicN | 0.632 | < .001 | 0.523 | 0.741 |
|  |  |  | RtoS | 0.599 | < .001 | 0.347 | 0.850 |
|   |   |   | RwLtoS | 0.487 | < .001 | 0.337 | 0.637 |

Table S12. *Summary of the standardized regression coefficients between the latent variables of utterance fluency and their 95% confidence intervals in the final SEM model*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|   |   |   |   |   |   | 95%CI |
| Latent variable | Direction | Latent variable | Task | *β* | *p* | Lower | Upper |
| ***Covariance between latent variables*** |  |  |  |  |  |
| Speed fluency | vs. | Breakdown fluency | Arg | -0.818 | < .001 | -0.951 | -0.686 |
|  |  |  | PicN | -0.822 | < .001 | -0.918 | -0.726 |
|  |  |  | RtoS | -0.800 | < .001 | -0.919 | -0.681 |
|  |  |  | RwLtoS | -0.769 | < .001 | -0.876 | -0.662 |
| Speed fluency | vs. | Repair fluency | Arg | -0.749 | < .001 | -0.890 | -0.608 |
|  |  |  | PicN | -0.720 | < .001 | -0.858 | -0.583 |
|  |  |  | RtoS | -0.720 | < .001 | -0.899 | -0.540 |
|  |  |  | RwLtoS | -0.739 | < .001 | -0.900 | -0.578 |
| Breakdown fluency | vs. | Repair fluency | Arg | 0.864 | < .001 | 0.735 | 1.003 |
|  |  |  | PicN | 0.639 | < .001 | 0.496 | 0.782 |
|  |  |  | RtoS | 0.796 | < .001 | 0.642 | 0.950 |
|   |   |   | RwLtoS | 0.716 | < .001 | 0.553 | 0.879 |

*Note.* For the sake of interpretability of the direction of relationship between the latent variables of UF, the regression coefficients in Table S12 were computed without the inversion of the observed variables of breakdown fluency and repair fluency measures.

# Reference

Clark, H., & Fox Tree, J. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*(1), 73–111. https://doi.org/10.1016/S0010-0277(02)00017-3

De Jong, N. H. (2016). Predicting pauses in L1 and L2 speech: The effects of utterance boundaries and word frequency. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 113–132. https://doi.org/10.1515/iral-2016-9993

Götz, S. (2013). *Fluency in native and nonnative English speech*. John Benjamins.

JACET. (2003). *JACET List of 8000 Basic Words*. JACET English Vocabulary SIG.

Kormos, J. (2000). The timing of self-repairs in second language speech production. *Studies in Second Language Acquisition*, *22*(2), 145–167. https://doi.org/10.1017/S0272263100002011

Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum Associates.

Leonard, K. R., & Shea, C. E. (2017). L2 speaking development during study abroad: Fluency, accuracy, complexity, and underlying cognitive factors. *The Modern Language Journal*, *101*(1), 179–193. https://doi.org/10.1111/modl.12382

Raykov, T., & Marcoulides, G. (2006). *A first course in structural equation modeling* (2nd ed.). Lawrence Erlbaum Associates.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. Harper Collins.

Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT Journal*, *65*(1), 71–79. https://doi.org/10.1093/elt/ccq020

Williams, S. A., & Korko, M. (2019). Pause behavior within reformulations and the proficiency level of second language learners of English. *Applied Psycholinguistics*, *40*(3), 723–742. https://doi.org/10.1017/S0142716418000802