**Supplementary Materials B: Codes for Analysis**

List of 79 coded features

**Publication information**

1. *Study number*
2. *Authors*
3. *Year of publication*
4. *Title*
5. *Journal*
6. *L1* (κ = 1, 100% agreement) - What was the L1 of the participants? (1 = English, 2 = Arabic, 3 = Unspecified (Belgium), 4 = Dutch, 5 = Unspecified (South-East Asia), 6 = Unspecified, 7 = French, 8 = Korean, 9 = Hebrew, 10 = Russian, 11 = Japanese, 12 = Chinese, 13 = Spanish, 14 = Portuguese, 15 = Italian, 16 = Persian, 17 = Polish, 18 = Turkish, 19 = Taiwanese, 20 = Vietnemese, 21 = Other (non-Romance), 22 = Cantonese, 23 = Mandarin, 24 = Slovak, 25 = Kazakhstani, 26 = Nepalese, 27 = Ukrainian, 28 = Bengali, 29 = Bulgaria, 30 = Farsi, 31 = Hindi, 32 = Kurdish, 33 = Luxemburgish, 34 = Malay, 35 = Romanian, 36 = Swedish, 37 = Swiss, 38 = Thai, 39 = Yoruba, 40 = Khmer, 41 = Javanese, 42 = Mauritian, 43 = Sinhalese, 44 = Tagalog) (omitted from final analysis)
7. *L2* (κ = 1, 100% agreement) - Which L2 was investigated? (1 = English, 2 = Spanish, 3 = German, 4 = Chinese, 5 = Turkish, 6 = Korean, 7 = French, 8 = Swahili, 9 = Italian) (omitted from final analysis)
8. *Hypothesis/RQ* (κ = 1, 100% agreement)- The hypothesis or research question that data on that particular line of coding related to.
9. *Treatment and schedule* (κ = 1, 100% agreement) - When, in terms of week number, did the various processes of the experiment take place (e.g., pretest [week 1], treatment [weeks 3 to 6], immediate posttest [week 6], delayed posttest [week 12]) (omitted from final analysis)
10. *Experiment length* (κ = 1, 100% agreement) - Duration of the entire experiment (1 = 1 week or less, 2 = 8 days to 2 weeks, 3 = 15 days to 3 weeks etc..., 98 = Unspecified, 99 = Mixed) (omitted from final analysis)
11. *ESL/EFL* (κ = 1, 100% agreement) - Learning context (1 = English as a second language [ESL], 2 = English as a foreign language [EFL])
12. *Age group* (κ = 1, 100% agreement) - Age of participants (0 = Unspecified, 1 = Elementary, Primary (under 12), 2 = Junior high school [JHS], high school [HS], Secondary school), 3 = University (including MA, MSc, PhD, etc...), 4 = Adult, 5 = Elementary & University comparison, 6 = Preschool)
13. *n1* (κ = .99, 99.70% agreement) - Number of participants in control/comparison group
14. *n2* (κ = 1, 100% agreement) - Number of participants in treatment group
15. *Comparison vs. Control* (κ = 1, 100% agreement) - The two conditions, treatments, group, or sets of results being compared on that specific line of coding vs. true control compared to treatment
16. *Effect categorization* (κ = .86, 90.91% agreement) - Research design type (1 = Between group, treatment vs control [B-TCtrl], 2 = Between group, treatment vs comparison [B-TCom], 3 = Within group, pretest vs posttest [W-PP], 4 = Within group, one-sample compared with zero [W-O], 5 = Switching replications [incorporated into 7, W-CB], 6 = Within group, treatment vs comparison [W-TCom], 7 = Within group, counterbalanced [W-CB])
17. *Between/Within group* (κ = 1, 100% agreement) *-* (1 = Between-group comparison, 2 = Within-group comparison)
18. *Measurement* (κ = 1, 100% agreement) - Type of test or instrument used.

*Effects calculated from n, M, SD or test statistics when M/SD was omitted*

1. *M1* (κ = 1, 100% agreement) - Mean score of participants in control/comparison group
2. *SD1* (κ = 1, 100% agreement) - Standard deviation of *M1*
3. *M2* (κ = .94, 94.11% agreement) - Mean score of participants in treatment group
4. *SD2* (κ = .94, 94.11% agreement) - Standard deviations of *M2*
5. Test statistic(s) of control/comparison group (κ = 1, 100% agreement)
6. Test statistic(s) of treatment group (κ = 1, 100% agreement)