**Supplementary material 1.** Description of the moderators

|  |  |  |
| --- | --- | --- |
|  | **Codes** | **Description** |
| 1. Publication year | [many; e.g., 2017] | the year that a paper is officially published |
| 2. Publication type | (1) SSCI/ESCI  (2) General journals  (3) Others | SSCI/ESCI: journal articles indexed in the Web of Science as SSCI/ESCI journals  General journals: journal articles that are not published in the SSCI/ESCI journals  Others: e.g., conference papers, dissertations/theses |
| 3. Participant age | (1) under 18  (2) 18 and above |  |
| 4. Participant proficiency | (1) Beginner  (2) Intermediate  (3) Advanced  (4) Mixed  (5) Not given | defined by the standardized tests or given by the teachers/researchers in the primary studies |
| 5. ASR feedback | (1) Explicit corrective  (2) Indirect | Explicit corrective: if the tool provides comprehensive feedback on the participants’ speech (e.g., comparing the participants’ speech input with the native speaker’s in terms of the speech waveform and its spectrogram and/or delivering feedback messages on the pronunciation production)  Indirect: if the tool only transcribes what the participants say and put these into text; or only gives positive response (e.g., correct, yes) and negative response (e.g., incorrect, no) to participants’ voice input |
| 6. Platform | (1) Computer  (2) Mobile | Computer: if the ASR application is built on computer devices (e.g., laptop, desktop)  Mobile: if the ASR application is built on mobile devices (e.g., smartphone, tablets) |
| 7. Target measure | (1) Segmental  (2) Suprasegmental  (3) Both | Segmental: measure of segmental features of pronunciation (e.g., sound units in isolation)  Suprasegmental: measure of suprasegmental features of pronunciation (e.g., intonation, stress, linking, rhythm, etc.) |
| 8. Assessment type | (1) Controlled  (2) Free | Controlled: if the assessment items require fixed response from all participants (e.g., read aloud given texts, pronounce given words)  Free: if the assessment items allow a variety of different responses from participants (e.g., open-ended measures, conversation, presentation) |
| 9. Treatment duration | (1) Short  (2) Medium  (3) Long | Short: 1–4 weeks  Medium: 5–8 weeks  Long: 9 weeks and above |
| 10. Learning activity | (1) alone  (2) with peer  (3) with teacher  (4) mixed | Participants mainly practice pronunciation with an ASR tool either “alone,” “with peers” or “with teacher”; “mixed” means students mainly practice with both teachers and peers and/or alone. |
| 11. Setting | (1) inside the classroom  (2) outside the classroom  mixed |  |

**Supplementary material 2.** Figure 1. PRISMA Flowchart (Moher et al., 2009)

Records identified through database searching: k = 7611

Records identified through reference chasing: k = 1614

Records screened

k = 448

Full-text articles assessed for eligibility: k = 38

Full-text articles excluded

k = 23

No control group: 13

Not enough information for Hedges’s *g* calculation: 2

Implementing more technological functions other than ASR: 8

Studies included in meta-analysis: k = 15

Identification

Screening

Eligibility

Included