**Appendix 1: References of included studies**

Al-shehri, M. A. A. (2015). *Incidental EFL vocabulary learning from listening to stories : The effects of frequency of exposure , elaboration and learner factors*. Lancaster University, UK.

Alahirsh, H. (2014). *Exploring the effectiveness of extensive reading on incidental vocabulary acquisition by EFL learners: An experimental case study in a Libyan university*. (Unpublished doctoral dissertation). University of Nottingham, UK.

Daskalovska, N. (2016). Acquisition of three word knowledge aspects through reading. *Journal of Educational Research*, *109*(1), 68–80. https://doi.org/10.1080/00220671.2014.918530

Day, R. R., Omura, C., & Hiramatsu, M. (1992). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language*, *7*, 541–541.

Donkaewbua, S. (2009). Vocabulary learning through listening in another language. Saarbrücken, Germany: Lambert Academic Publishing.

Feng, Y., & Webb, S. (2019). Learning vocabulary through reading, listening, and viewing: Which mode of input is most effective? *Studies in Second Language Acquisition*, 1–25. https://doi.org/10.1017/S0272263119000494

Hatami, S. (2017). The differential impact of reading and listening on L2 incidental acquisition of different dimensions of word knowledge. *Reading in a Foreign Language*, *29*(1), 61–85.

Koolstra, C. M., & Beentjes, J. W. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development*, *47*(1), 51–60. https://doi.org/10.1007/BF02299476

Miyasako, N. (2002). Does text-glossing have any effects on incidental vocabulary learning through reading for Japanese senior high school students? *Language Education & Technology*, *39*, 1–20. https://doi.org/10.24539/let.39.0\_1

Neuman, S. B., & Koskinen, P. (1992). Captioned television as “comprehensible input”: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, *27*(1), 95–106. https://doi.org/10.2307/747835

Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, *41*(4), 745–768. https://doi.org/10.1017/S0272263119000020

Pellicer-Sánchez, A., & Serrano, R. (2017, March). *Incidental L2 vocabulary learning from reading and reading-while-listening*. Paper presented at American Association of Applied Linguistics (AAAL), Portland, US.

Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *40*(3), 551–577. https://doi.org/10.1017/S0272263117000407

Pitts, M., White, H., & Krashen, S. (1989). Acquiring second language vocabulary through reading: A replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language*, *5*(2), 271-275.

Reynolds, B. L. (2015). A mixed‐methods approach to investigating first‐and second‐language incidental vocabulary acquisition through the reading of fiction. *Reading Research Quarterly*, *50*(1), 111–127. https://doi.org/10.1002/rrq.88

Reynolds, B. L. (2018). The effects of nonce words, frequency, contextual richness, and L2 vocabulary knowledge on the incidental acquisition of vocabulary through reading: More than a replication of Zahar et al. (2001) & Tekmen and Dalog̈lu (2006). *IRAL - International Review of Applied Linguistics in Language Teaching*, 1–28. https://doi.org/https://doi.org/10.1515/iral-2015-0115

Rodgers, M. P., & Webb, S. (2019). Incidental vocabulary learning through viewing television. *ITL-International Journal of Applied Linguistics*. https//doi.org/10.1075/itl.18034.rod

Rott, S. (1999). The effect of exposure to frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Acquisition*, *21*(4), 589–619. https://doi.org/10.1017/S0272263199004039

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, *61*(1), 219–258. https://doi.org/10.1111/j.1467-9922.2010.00593.x

Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition*, *19*, 287-307.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, *28*(1), 46–65. https://doi.org/10.1093/applin/aml048

Webb, S., & Chang, A. (2015). Second language vocabulary learning through extensive reading with audio support: How do frequency and distribution of occurrence affect learning? *Language Teaching Research*, *19*(6), 667–686. https://doi.org/10.1177/1362168814559800

**Appendix 2: Examination of Publication Bias**

 Due to the fact that studies reporting statistically significant findings or large effect sizes tend to be published, we have a concern about overestimation of the aggregated effect sizes resulting from the biased samples of studies (Lipsey & Wilson, 2001). To assess publication bias in our data, we conducted the following two measures: fail-safe *N* and the trim-and-fill. First, the classic fail-safe *N* (Rosenthal, 1979) calculates how many missing studies with small effect sizes (effect size = zero) would be needed before the *p*-value for the observed mean effect became nonsignificant. The results show that we would need to retrieve 3,726 studies for the first posttest and 210 studies for the follow-up posttest to nullify the effect (alpha = .05), indicating that there would be less reason for any serious concern about publication bias. Second, we computed the trim-and-fill analysis (Duval & Tweedie, 2000a, 2000b) with the random-effects model that computes an unbiased estimate of the mean effect size by accounting for the number of missing studies in the current data set. The results show that the initial and adjusted values were not hugely different in the first posttest: *g =* 1.14, 95% CI [0.86, 1.41] => 1.22, 95% CI [0.94, 1.50] or in the follow-up posttest: *g* = 0.93, 95% CI [0.44, 1.42] => 1.11, 95% CI [0.64, 1.57], indicating no serious effect of publication bias on the current result. These findings from the two bias-detective measures suggest that there is less concern about publication bias in the current meta-analysis.

**References**

Duval, S., & Tweedie, R. (2000a). A nonparametric trim and fill method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*(449), 89–98. https://doi.org/10.1080/01621459.2000.10473905

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638