Psychonomic Bulletin & Review (2023) 30:897–913
https://doi.org/10.3758/s13423-022-02136-7

THEORETICAL/REVIEW

Does bilingualism come with linguistic costs? A meta-analytic review of the bilingual lexical deficit

Emanuel Bylund[1,2] · Jan Antfolk[3] · Niclas Abrahamsson[2] · Anne Marte Haug Olstad[4] · Gunnar Norrman[2] · Minna Lehtonen[4,5]

Accepted: 3 June 2022 / Published online: 3 November 2022

Abstract

A series of recent studies have shown that the once-assumed cognitive advantage of bilingualism finds little support in the evidence available to date. Surprisingly, however, the view that bilingualism incurs linguistic costs (the so-called lexical deficit) has not yet been subjected to the same degree of scrutiny, despite its centrality for our understanding of the human capacity for language. The current study implemented a comprehensive meta-analysis to address this gap. By analyzing 478 effect sizes from 130 studies on expressive vocabulary, we found that observed lexical deficits could not be attributed to bilingualism: Simultaneous bilinguals (who acquired both languages from birth) did not exhibit any lexical deficit, nor did sequential bilinguals (who acquired one language from birth and a second language after that) when tested in their mother tongue. Instead, systematic evidence for a lexical deficit was found among sequential bilinguals when tested in their second language, and more so for late than for early second language learners. This result suggests that a lexical deficit may be a phenomenon of second language acquisition rather than bilingualism per se.

Keywords Age of acquisition · Bilingualism · Lexical deficit · Executive control · Vocabulary

Does bilingualism cause any changes to the human mind? Up until recently, the standard answer to this question would have been that acquiring, mastering, and using two (or more) languages affords certain cognitive advantages, while also incurring certain linguistic costs. The advantages would be particularly salient in bilingual individuals' enhanced cognitive flexibility and ability to monitor and control their actions and impulses (i.e., executive functioning). The costs would be manifested as a so-called lexical deficit, a term implying that bilinguals, relative to monolinguals, exhibit smaller vocabulary knowledge and slower word retrieval. For instance, in a picture-naming task, bilingual participants would be more likely to take longer to retrieve the words for the displayed items and more likely to name fewer items, than their monolingual counterparts. Similarly, in a verbal fluency task, bilinguals would be likely to name fewer category members (e.g., all the vegetables they can think of, or all the words beginning with the letter "p") relative to monolingual participants.[1]

Recently, doubt has been cast upon the view that bilingualism affords a cognitive advantage. A number of studies have drawn attention to publication bias and inconsistencies in the findings on the bilingual advantage (e.g., de Bruin et al., 2015; Hilchey & Klein, 2011; Paap & Greenberg, 2013), and various large-scale investigations (e.g., Dick et al., 2019; Nichols et al., 2020) and comprehensive meta-analyses (Anderson et al., 2020; Donnelly et al.,

✉ Emanuel Bylund
mbylund@sun.ac.za

1 Department of General Linguistics, Stellenbosch University, Stellenbosch, South Africa

2 Stockholm University, Stockholm, Sweden

3 Åbo Akademi University, Turku, Finland

4 University of Oslo, Oslo, Norway

5 University of Turku, Turku, Finland

[1] It is customary to point out that even if bilingual speakers would know fewer words in each of their languages than monolingual speakers, the total number of words they know (i.e., the vocabularies of their two languages combined) is arguably larger than that found in monolingual speakers. Moreover, the bilingual lexical deficit may be more clearly manifested in picture naming than verbal fluency tasks. This is because the latter presumably relies more on extralinguistic processes, including executive control abilities related to strategic retrieval of different words that meet certain constraints (Shao et al., 2014).

2019; Lehtonen et al., 2018) have found either a small or no effect of bilingualism on cognitive functioning.

Surprisingly, the view that bilingualism entails a lexical deficit has, as of yet, not been subjected to nearly as much scrutiny. The current study addresses this gap by means of a comprehensive meta-analysis of the evidence available to date. Testing the extent and robustness of any detected lexical deficit is important for a number of reasons: First, at a more general level, knowledge about the potential effects of bilingualism on linguistic abilities is crucial for advancing our understanding of the human capacity for language (Bialystok & Werker, 2017). Second, recent findings show that certain verbal behaviours that have been characterized as a "cost" of bilingualism (e.g., diverging knowledge in phonology and grammar) are in fact a consequence of language learning history (e.g., age of language acquisition) and not of bilingualism per se (Bylund et al., 2019; Bylund et al., 2021; Hyltenstam et al., 2009; Norrman & Bylund, 2016; Veríssimo et al., 2018). Whether the same applies to the lexical deficit remains to be seen.

Looking closer at the bilingual lexical deficit also has relevance for evaluating current models on language development. Assuming principles of Hebbian learning at its core, one set of models predict that *amount of exposure* to a language is the driving mechanism behind word learning: In order for a given lexical item to be acquired and activated, a certain amount of exposure is necessary (e.g., Gollan et al., 2005b; Gollan et al., 2011). Because bilinguals have to divide the hours of the day between their two languages, it is impossible for them to receive the same amount of exposure to each language as monolingual individuals do to their one language. As a consequence, lexical representations in bilingualism are weaker and have higher activation thresholds. Another set of frameworks instead holds that it is the *timing of exposure* that primarily determines the strength of lexical representation (e.g., Hernandez et al., 2005; P. Li, 2009; Werker & Hensch, 2015). While recognizing the importance of exposure, these approaches hold that languages learnt after the mother tongue (L1)—even if learnt in early childhood—will be parasitic on L1 representations and, moreover, exhibit nonnativelike phonetic categories, which will ultimately compromise lexical development (Choi et al., 2018; Rivera-Gaxiola et al., 2005).

These approaches to lexical development differ with regards to their predictions of the existence of a lexical deficit. Whereas the amount-of-exposure models regard the lexical deficit as an integral part of bilingualism, the timing-of-exposure models, though not necessarily questioning the existence of a lexical deficit, are compatible with the idea that bilingualism per se need not incur linguistic costs: if both languages are acquired from birth, there will be no prior entrenched system that leads to lexical parasitism, and no compromised phonetic categories.

The following three alternative hypotheses may be formulated:

Hypothesis 1: There is a lexical deficit associated with bilingualism.

This is the default hypothesis. Its confirmation would be consistent with the notion that bilingualism, as predicted by the amount-of-exposure accounts, inevitably leads to reduced linguistic ability in the area of the lexicon.

Hypothesis 2: There is a lexical deficit, but it depends on language learning history.

This hypothesis posits that bilinguals who acquire both languages from birth are less likely to exhibit a lexical deficit than those who become bilingual later in life. This prediction is consistent with the timing-of-exposure frameworks.

Hypothesis 3: There is no lexical deficit associated with bilingualism.

This hypothesis suggests that there is no systematic evidence, across studies, of a bilingual lexical deficit. To the extent that the analyzed studies include bilinguals with different learning histories, such an outcome would challenge both the amount-of-exposure and timing-of-exposure accounts.

These hypotheses speak directly to potential moderators of the lexical deficit. The first moderator is the amount of exposure that the bilingual receives in their languages. Under the assumption that input received is a major determinant of language development, this moderator should show that more exposure to a language is associated with a smaller degree of lexical deficit in that language. The second moderator is bilingualism type. Here, a distinction is made between two different types of bilinguals: (a) "simultaneous bilinguals"—that is, individuals who acquire both languages from birth—and (b) "sequential bilinguals"—that is, individuals who acquire one language from birth (i.e., the L1), and a second language (L2) after that. A third moderator, which also relates to language learning history, is age of L2 acquisition—that is, the age at which the L2 was acquired. Even though sequential bilinguals have in common that they acquired their two languages in sequence, they may differ with respect to the specific age at which the L2 was learnt. Similar to findings on other domains of language, such as syntax and phonology (e.g., Abrahamsson, 2012; Abrahamsson & Hyltenstam, 2008; Flege et al., 1999; Granena & Long, 2013; Johnson & Newport, 1989), age of L2 acquisition may be inversely correlated with nativelike lexical behaviour in the L2.

Method

The study, including the statistical procedures, was preregistered on AsPredicted (#52642, https://aspredicted.org/VF7_8SR).

Literature searches

We searched the electronic databases PsycINFO, PubMed, Web of Science, and Google Scholar by combining terms related to participant groups, such as "bilingual" and "monolingual," and terms related to vocabulary and lexical processing and the common task paradigms to measure expressive vocabulary. Expressive vocabulary was chosen because production (as opposed to comprehension) requires higher activation thresholds of linguistic structures (e.g., Paradis, 2004). Therefore, the deficit should be most visible in production. As a number of studies that primarily focus on the bilingual cognitive advantage also report measures of expressive vocabulary from their samples, we included such terms (e.g., "executive function") in the search strings. We first tested the sensitivity of our search strings by checking how many of 20 preidentified relevant papers would be found in the search. These papers either had the bilingual lexicon as a primary focus, or primarily concerned cognitive aspects of bilingualism and reported lexical behaviour as a bilingual background measure (for a list of papers and exact search strings, see Supplemental Table S1). Because all these papers were found, we deemed that the search strings were together sufficiently sensitive. The various stages of the search process are detailed in Fig. 1. In addition to the database searches, we screened the reference lists of 20% of the already identified and included articles for additional potentially relevant papers. The first search was conducted in February 2020, and the second search was conducted in June 2021. In case critical measures could not be obtained from a study that had been deemed relevant, the authors were contacted via email. The response rate to our emails was 59.5%, and out of these, 48.8% resulted in new data (Supplemental Table S2 lists the authors who were able to assist us with additional data).

Inclusion and exclusion criteria

In order for a study to be included, it had to be an original study including a comparison of bilingual and monolingual participants in at least one measure of expressive vocabulary or lexical production (oral or written). We excluded studies that only reported the results of a bilingual sample because comparison with a monolingual sample is needed to measure lexical deficits related to bilingualism. We only included published peer-reviewed journal articles written in the English language.

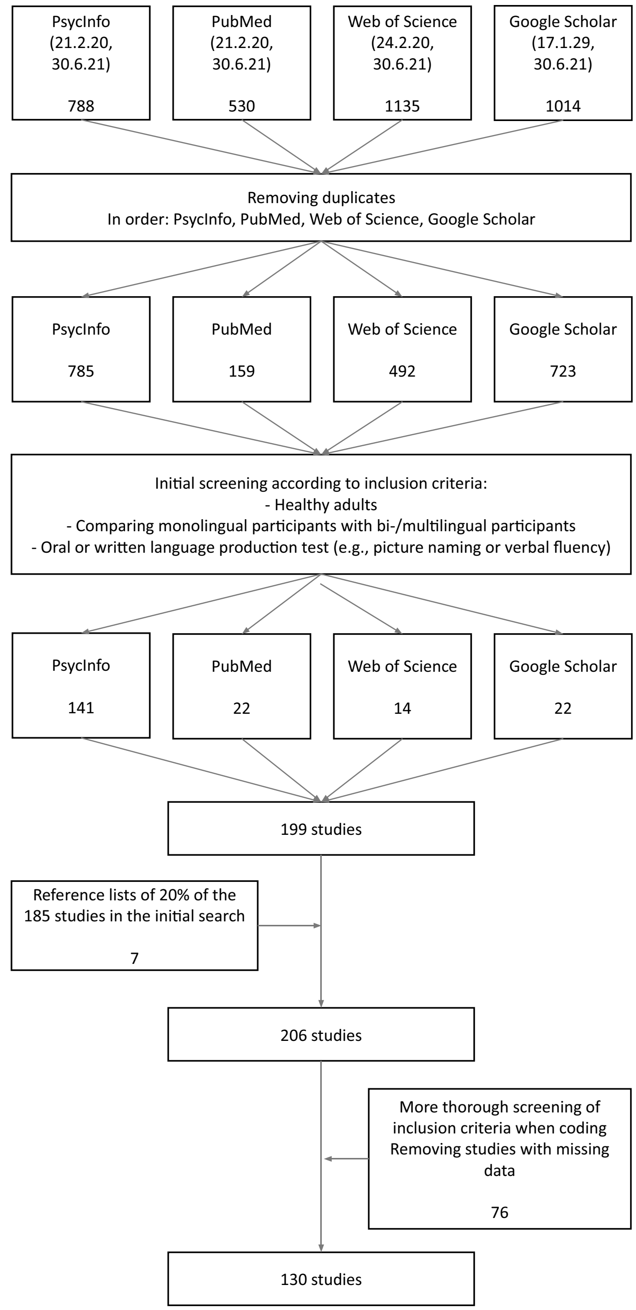

Fig. 1 The stages of the literature search and screening process

Inclusion related to participants We only included samples of adult participants (minimum mean age of the sample 18 years). We included hearing and deaf participants but excluded samples of participants with other relevant disabilities or neurocognitive impairments and diseases (e.g., dementia or aphasia), case studies, and participants undergoing brain surgery. Due to large variation in the definitions of bilingualism used in the field, we initially relied on the original study's definitions of participants as bilingual and monolingual and coded the relevant sample characteristics (such as age of acquisition and proficiency) for more detailed moderator analyses. Samples with "novice" L2 learners were excluded based on the criterion that the bilingual participants in the original study needed to have had, on average, at least 5 years of experience of a L2 in order to be considered functionally bilingual (Johnson & Newport, 1989). Bidialectals were excluded because dialects commonly share much of the vocabulary in a language. Bimodal bilinguals were not excluded.

Inclusion related to tasks and measures We included measures of lexical production, such as picture naming, verbal fluency, and synonym and antonym production. Observational and interview studies and studies only reporting physiological responses, such as brain imaging results, were excluded. As the focus of the study was on lexical production, we excluded narrative production or picture description tasks, tasks with a metalinguistic dimension (e.g., word definitions), oral repetition and reading tasks, and tasks involving morphosyntactic proficiency (e.g., cloze tests). We excluded measures from spontaneous speech and tasks that involve switching between languages. We also excluded word production tasks that are assumed to also measure other cognitive functions such as creativity and cognitive flexibility.

Data coding

From the included studies, we coded the following variables:

- *Study characteristics*: Year of publication, country in which the study was conducted.
- *Participant characteristics*: Sample number (i.e., monolingual-bilingual group comparison, running number across original studies), mean age of each group, type of bilinguals (simultaneous, sequential), age of L2 acquisition, L1 of each group, bilinguals' L2, bilinguals' possible additional languages, reported exposure to the L2.[2] Type of bilingual was determined first through how the groups were described in the paper; where relevant descriptions were lacking, groups with a mean age of acquisition above 1.5 years were coded as sequential. However, if any participants in such a group were reported as having acquired their languages from birth, the group was coded as "mixed" (because it contained both simultaneous and sequential bilinguals). When sufficient information was not provided the groups were coded as "undefined."
- *Matching of groups*: Whether the bilingual and monolingual groups were matched for IQ and whether they were matched for education. In cases when only student populations were included, these were coded as education matched.
- *Task characteristics*: The task used, task type (letter/category fluency, picture naming), language of the task, modality of the task (oral, written).
- *Measures*: Whether the measure is based on reaction time or accuracy. *N,* mean and standard deviation for the measure in each group. Whether the same group is used in multiple comparisons.

Interrater reliability

The studies were coded by two raters who first coded a number of papers (e.g., 10) and then checked that their coding was uniform and resolved discrepancies through discussion. The majority of studies were then coded by a single coder. Finally, one fifth of all studies were coded by a new coder and the interrater reliability (Cohen's kappa) was calculated. As per the pre-registration, we considered a Cohen's kappa value of at least 0.75 as acceptable reliability. All kappa values were found to be at 0.8 or higher for key variables (see Supplemental Table S3). All discrepancies were resolved by discussion.

Statistical analyses

We used the package metafor (Viechtbauer, 2010) for R (Version 4.0.3. for Mac; R Core Team, 2017) to conduct statistical analyses. The R script for the reported analyses, output files, and the data file can be accessed at the project's domain at the Open Science Framework (https://osf.io/8tepa/).

We first calculated effect sizes based on the available data for bilinguals and monolinguals. Next, we evaluated the data structure and assessed the distribution of data before proceeding with synthesizing effect sizes. We further investigated how synthesized effects were affected by moderators and biases in the distribution of effect sizes.

Calculating effect sizes To obtain effect sizes for the differences in monolinguals' and bilinguals' performances, we calculated the standardized mean difference (SMD) using the escalc function, which produces a Hedges's *g* by adjusting the positive bias. Hedges's *g* is a close equivalent to Cohen's *d,* but is less biased when sample sizes are small, and is therefore often considered a corrected effect size (e.g., Borenstein et al., 2011). We set the v-type argument to "UB" for unbiased estimates of the sampling variance (Viechtbauer, 2010).

In most tasks, a higher value (e.g., score, percentage correct responses) indicated a better performance. In some

[2] We included studies where exposure was reported as (or could be converted into) a percentage, but not studies reporting exposure as years or a score.

Table 1 Model fit indices, model comparison statistics, and variance components

Levels	Added levels	Model fit indices		Model comparison		Variance components		
		AIC	LogLIK	Models	LRT	σ^2_1	σ^2_2	σ^2_3
1. One		2343.17	−1170.58					
2. Two	Test variants	818.36	−407.28	1 vs. 2	1526.80***	0.38		
3. Three	Tests	758.36	−376.18	2 vs. 3	62.01***	0.38	0.05	
4. Four	Exp. Groups	747.29	−369.65	3 vs. 4	13.07***	0.19	0.18	0.05

AIC = Akaike information criterion; LogLik = Log-Likelihood; LRT = Likelihood-Ratio Test. The Likelihood-Ratio Test statistic is tested against a chi-square distribution with 1 degree of freedom. ***$p < .001$

tasks, a lower value (e.g., percentage incorrect, reaction times) indicated a better performance. In the latter case, we reversed effect sizes by multiplying them by −1, which allowed us to interpret all positive effect sizes as corresponding to an increased lexical deficit (i.e., monolinguals outperforming bilinguals on lexical tasks) and negative effect sizes as bilinguals outperforming monolinguals on lexical tasks.

Multilevel modelling We compared the model fit indices of the one-, two-, three-, and four-level models using a likelihood-ratio test. To do this, we used the anova.rma function in metafor (Viechtbauer, 2010). Suggesting that our four-level model best fitted the data structure, all tests were statistically significant (Table 1). To evaluate the multi-level model, we used data trimmed for outliers (see Data Screening and Assessment of Bias under Results).

The within-comparison dependency can be represented by an intraclass correlation coefficient (ICC), which is calculated by dividing the variance between comparisons by the sum of the variance between and within comparisons (i.e., $\sigma_1^2/(\sigma_1^2+\sigma_2^2)$). If the variance within comparisons is small compared to the variance between comparisons, the ICC value is high. If, for instance, outcomes from different test variants co-vary so much that they provide the same information from the same test, the ICC will equal 1. Conversely, if they contribute completely independent information (i.e., they are no more similar than are two outcomes of two different tests), the ICC drops towards zero. In our final four-level model, the ICC for test variants within tests was .506, and the ICC for tests within comparisons of experimental groups was .789.

Assessment and correction of bias We first visually inspected the data using funnel plots. In meta-analyses, funnel plots are often used to assess bias in data dispersion that might be due to publication bias or any other cause of systematic differences in outcomes between studies with low and high precision. In a typical funnel plot, effect sizes are plotted on the *x*-axis (from negative to positive, left to right) and the precision/variance on the *y*-axis (with more robust estimates higher up). Assuming that the true dispersion of effect sizes only stems from sampling bias and therefore should be symmetric, the funnel plot can be used to assess whether effect sizes are asymmetric and missing in specific areas of the plot (e.g., null effects from small studies are systematically underreported; Palmer et al., 2008). Here, we used contour-enhanced funnel plots (Peters et al., 2008), where the assessment is facilitated by shades describing statistical significance (at different levels of α) that are superimposed on the funnel plot. We plotted effect sizes against both the precision (1/*SE*) and the standard error (*SE*). Contours were added at three levels of α ($p = .10$, $p = .05$, and $p = .005$), and with a vertical reference line at Hedges's $g = 0$.

To statistically assess and adjust for observed asymmetries in the distribution of effect sizes, we conducted a PET-PEESE analysis (Stanley & Doucouliagos, 2014). In the precision-effect test (PET) and the precision-effect test with standard error (PEESE), outcomes are regressed on their variance (SE^2) or *SE*, respectively, in weighted least-squares regressions. In case there is a systematic bias such that studies with low precision over- or underestimate effect sizes, this bias is represented by an association between *SE* and the size and direction of the effect sizes. Modelling studies have shown that the intercept in the PEESE regression underestimates the true effect (in this case, the outcome of a hypothetical study with perfect precision). Entering the SE^2 instead of *SE* as a predictor in a PET model provides a more accurate estimate of a true effect.

Moderator analyses We analyzed the following variables as moderators of the bilingual lexical deficit: bilingualism type (i.e., simultaneous or sequential bilingualism), amount of language exposure, and age of L2 acquisition. Based on previous research on the acquisition of L2 lexis, we used age 10 as a divider between early and late acquirers (Granena & Long, 2013). By means of exploratory analyses, we further examined the moderating influence of testing language (i.e., whether bilinguals were tested in their L1 or L2), task type, and measure type (i.e., time or accuracy based).

Results

Descriptive results

The dataset included 478 effect sizes from 130 separate studies. Of the 478 effect sizes, 120 represented category fluency, 136 letter fluency, and 222 picture-naming tests. The bilingual groups consisted of simultaneous (5.02%), sequential (45.61%), mixed (31.38%), and undefined bilinguals (17.99%). See Supplemental Table S4 for more descriptive results.

Data screening and assessment of bias

We firstly examined the distribution of effect sizes and their precision to assess asymmetries in the distribution of effect sizes such that strong positive or negative effects were overrepresented in studies with low precision, due to, for example, publication bias. To do this, we produced contour-enhanced funnel plots (Fig. 2) representing the distribution of effect sizes within each of the four types of bilingual groups. For each group, we plotted the effect sizes both against the inverse standard error (1/*SE*) and the standard error (*SE*) and on the *y*-axis. Visual inspection of these plots revealed an asymmetry in the distribution of effect sizes from comparisons including sequential and mixed bilinguals, where studies with low precision and weak or negative effects (i.e., evidence against a lexical deficit) were relatively absent. In comparisons including undefined bilinguals, there was no clear sign of asymmetry, and in comparisons including simultaneous bilinguals, one study with low precision produced a large negative effect.

The visual inspection also revealed several unrealistically large effect sizes (some of these are outside the plotted area) and studies with uncharacteristically low precision. To reduce asymmetry, increase precision in subsequent analyses, and decrease the influence of outliers on the PET-PEESE analyses, we trimmed the data by removing observations ($k = 15$; see Supplemental Table S5 for the removed effect sizes) with a variance above 0.55 and effect sizes larger than $g = [3.00]$, as such observations were clearly outside the typical data distribution (which, for example, could indicate errors in the original studies).

Lexical deficit

After trimming the data, we investigated the lexical deficit using all comparisons irrespective of the type of bilinguals used in the comparison. We found a medium effect size, $g = 0.52$ [0.42, 0.62], $p < .001$, $k = 463$, showing that monolinguals performed better than bilinguals. The test for heterogeneity was significant, $Q[462] = 2664.83$, $p < .001$. We then assessed the extent to which the observed effect was due to publication bias. Both the PET test and the PEESE test showed positive associations between the effect sizes and their *SE* and variance ($p = .025$ and $p < .001$, respectively). The adjusted effect size was smaller but remained positive and statistically significant, $g = 0.23$ [0.18, 0.27], $p < .001$.

To follow-up on this test, we conducted a sensitivity test to investigate whether the results would be different if we used all data (i.e., also including the removed outliers). With the outliers included, both the unadjusted effect size, $g = 0.54$ [0.42, 0.66], $p < .001$, $k = 478$, and the adjusted effect size, $g = 0.38$ [0.35, 0.40], $p < .001$, were slightly larger (see Fig. 3 for plots of the association between effect sizes and their variance).

Lexical deficit by type of bilingual group

Because there may be differences depending on the type of bilinguals that the monolinguals were compared with, we investigated whether the type of bilingual group moderated the outcome. First, we compared effect sizes for simultaneous and sequential bilinguals. We found no statistically significant difference, Q_M [1] = 1.89, $p = .169$, $k = 233$, and the residual heterogeneity remained significant, Q_E [231] = 1477.30, $p < .001$. The estimated average effect was positive for both groups. Crucially, however, the effects were statistically significant for the sequential, but not for the simultaneous bilinguals. The PET-PEESE adjusted effect sizes were smaller but remained positive (see Fig. 4).

After this, we also included the mixed and undefined bilinguals. Again, we found no statistically significant differences, Q_M [3] = 3.27, $p = .352$, $k = 463$. The estimates were positive also for mixed, $g = 0.51$ [0.33, 0.69], $p < .001$, $k = 144$, and undefined bilinguals, $g = 0.37$ [0.13, 0.6], $p = .002$, $k = 86$. After a PET-PEESE adjustment, estimates were slightly lower for mixed ($g = 0.39$ [0.17, 0.61], $p < .001$) and undefined bilinguals ($g = 0.26$ [−0.00, 0.53], $p = .051$).

Testing language[3]

We then explored whether the language used to test the bilingual group (L1 or L2) moderated the effects. We found a statistically significant difference, Q_M [1] = 96.34, $p < .001$, $k = 463$, showing that when tests were administered in L1, the lexical deficit was smaller, $g = 0.11$ [−0.01, 022], $p = .065$, $k = 199$, than when tests were administered in L2, *g*

[3] This moderator was not specified in the preregistration, and the analysis is thus exploratory.

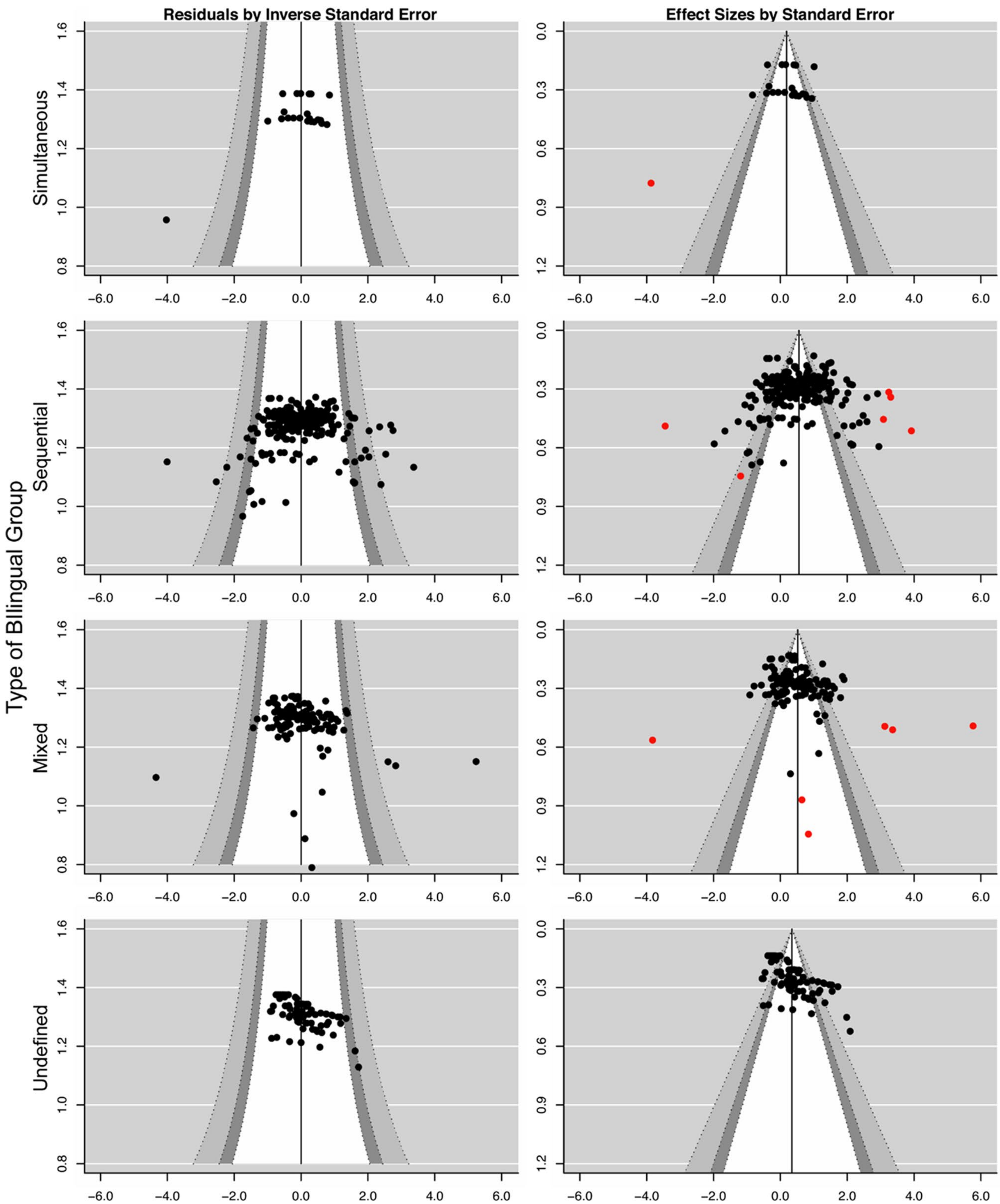

Fig. 2 Contour-enhanced funnel plots for each type of bilingual comparison group (by row). Shades (white, light gray, dark gray) represent $p = .10$, $p = .05.$, and $p = .005$, respectively. In the left column, residuals are plotted against the inverse standard error (*SE*) on the *y*-axis. The *x*-axis represents deviance from the average effect among the respective subgroup of bilingual comparison groups. In the right column, effect sizes are plotted against the standard error (*SE*). The *x*-axis represents the effect size, and effect sizes outside the range of $g = -3.00$ to $g = 3.00$ or with a variance above 0.55 are highlighted in red. Some effect sizes are outside the range of the plot

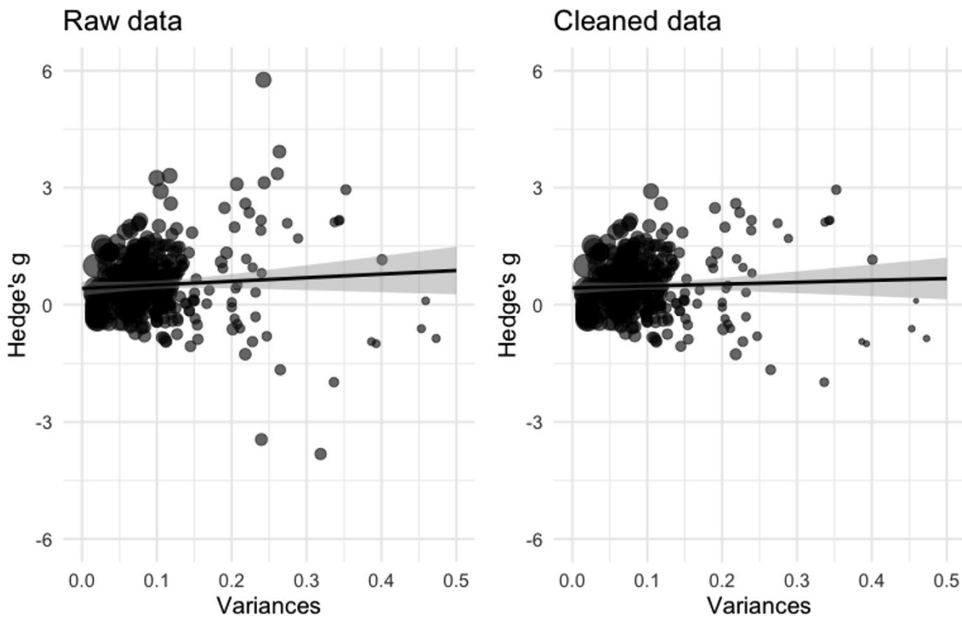

Fig. 3 Plots of the regression slope (black line) between the effect size (*y*-axis) and the variance (*x*-axis). The shaded area gives the 95% confidence interval. The effect size of a hypothetical study with perfect precision is estimated at variance = 0. The regression for raw data is shown in the left panel; the regression for trimmed data is shown in the right panel

= 0.86 [0.76, 0.97], $p < .001$, $k = 264$. The PET-PEESE adjusted values were $g = -0.05$ [−0.21, 0.12], $p = .58$ and $g = 0.73$ [0.58, 0.87], $p < .001$, respectively.

Including only sequential bilinguals in the analysis, we found an even larger difference $Q_M = [1] = 91.69$, $p < .001$, $k = 211$. When tests were administered in L1, there was no evidence of lexical deficit, $g = 0.01$ [−0.15, 0.16], $p = .927$, $k = 104$., but when tests were administered in L2, there was evidence of lexical deficit, $g = 1.06$ [0.91, 1.21], $p < .001$, $k = 107$. The pattern remained similar after adjusting for bias, $g = -0.08$ [−0.32, 0.15], $p = .499$ for L1, and $g = 0.99$ [0.79, 1.19], $p < .001$ for L2.

Age of L2 acquisition

Next, we investigated whether there was an association between age of acquisition and lexical deficit. Because current theories and evidence on age of acquisition concern L2 development, this analysis was circumscribed to the group of sequential bilinguals tested in their L2. We did this, firstly, by categorizing age of L2 acquisition according to the reported group mean (when this information was available) into two groups: before 10 years of age ($k = 68$) and after 10 years of age ($k = 32$). We found a statistically significant difference between the groups, $Q_M = [1] = 4.29$, $p = .038$, $k = 100$. As would be expected, in the group of participants acquiring a L2 before ten years of age, $g = 0.98$ [0.78, 1.17], $p < .001$, the synthesized effect was smaller than in those who had acquired the L2 after age 10, $g = 1.34$ [1.06, 1.63], $p < .001$. After a PET-PEESE adjustment, the effect sizes were smaller but remained positive, $g = 0.50$ [0.22, 0.78], $p = < .001$ and $g = 0.83$ [0.47, 1.18], $p < .001$.

Language exposure

We proceeded to investigate the association between language exposure and the lexical deficit. Because increased exposure to a given language is expected to lead to a lower lexical deficit in that language, but an increased deficit in the other language, we tested simultaneous and sequential bilinguals separately. The association between exposure to the testing language and lexical deficit failed to reach statistical significance in both simultaneous (Q_M [1] = 3.76, $p = .052$, $g = 0.02$ [−0.00, 0.05], $k = 12$) and sequential bilinguals ($Q_M = [1] = 2.20$, $p = .138$, $g = 0.01$ [−0.00, 0.01], $k = 102$). When further limiting the data to include only sequential bilinguals tested in L2, there was no association, Q_M [1] = 0.10, $p = .756$, $g = -0.00$ [−0.01, 0.01], $k = 56$.

Task and measure type[4]

Finally, we explored whether measure and task type moderated the effects. Measure and task type were strongly associated, $\chi^2[2] = 73.04$, $p < .001$, with 31.3% of the 211

[4] This moderator was not specified in the preregistration and the analysis is thus exploratory.

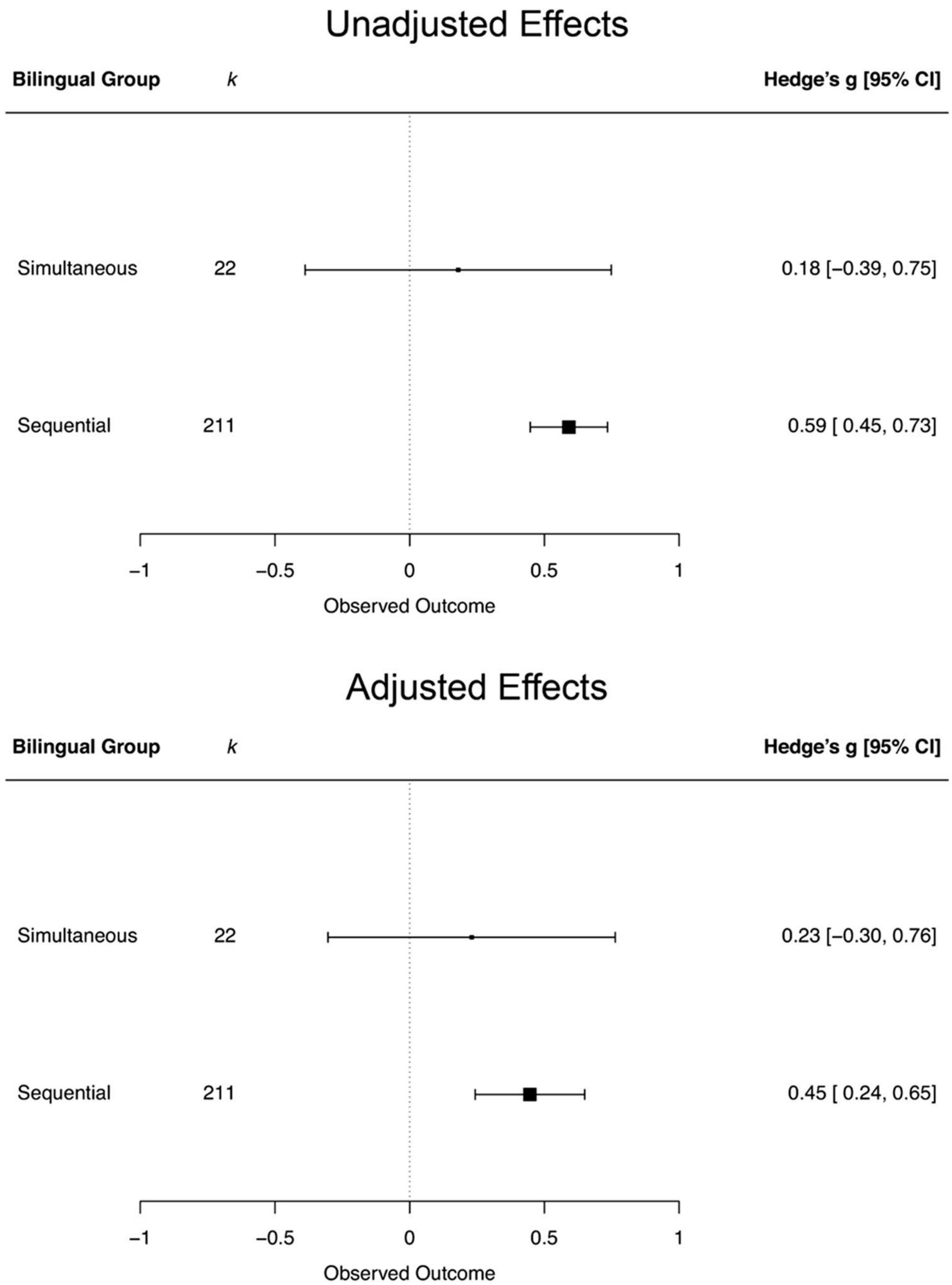

Fig. 4 For each type of bilingual group, the figure displays synthesized effect sizes and 95% confidence intervals (CI) for the comparison between monolinguals and bilinguals. Positive values indicate monolinguals perform better; negative values indicate bilinguals perform better. k = number of effect sizes. The upper panel displays unadjusted effects; the bottom panel displays adjusted effects

picture-naming measures being represented by reaction time, whereas only 2.5% of the 116 category-fluency measures and 2.3% of the 130 letter-fluency measures were reaction times.

We found a statistically significant difference between accuracy measures and reaction time measures, $Q_M = [1] = 12.36$, $p < .001$, $k = 463$. The synthetized effect was lower for accuracy measures, $g = 0.49$ [0.39, 0.59], $p < .001$, $k = 392$, than for reaction times, $g = 0.74$ [0.60, 0.88], $p < .001$, $k = 72$. The adjusted estimates were $g = 0.35$ [0.19, 0.52], $p < .001$ for accuracy and $g = 0.56$ [0.37, 0.75], $p < .001$ for reaction time measures.

After this, we inspected whether task type moderated the effects, and found a statistically significant difference, $Q_M = [2] = 114.46$, $p < .001$, $k = 463$. Letter fluency tasks produced the smallest effects, $g = 0.15$ [0.04, 0.27], $p = .008$, $k = 133$, followed by category fluency tasks, $g = 0.31$ [0.20, 0.43], $p < .001$, $k = 119$, and picture naming tasks, $g = 0.75$

[0.65, 0.85], $p < .001$, $k = 211$. After a PET-PEESE adjustment, the estimates were slightly lower. The estimate was not statistically significant for letter fluency, $g = 0.04$ [−0.12, 0.21], $p = .598$. For category fluency and picture naming, the estimates were $g = 0.21$ [0.05, 0.37], $p = .012$ and $g = 0.64$ [0.48, 0.79], $p < .001$, respectively.

Because measure type and task type strongly co-varied (and only picture-naming tasks included a sufficient number of both accuracy and reaction time measures), we compared accuracy and reaction time measures within picture-naming tasks. The difference between the two measure types was not statistically significant, Q_M [1] = 3.42, $p = .064$, $k = 211$. Reaction time measures $g = 0.89$ [0.74, 1.04], $p < .001$, $k = 66$, and accuracy measures, $g = 0.77$ [0.65, 0.89], $p < .001$, $k = 145$, produced similar effects. After adjusting for bias, the estimates were slightly smaller, $g = 0.54$ [0.32, 0.77], $p < .001$ for reaction times and $g = 0.44$ [0.24, 0.64], $p < .001$ for accuracy.

Discussion

The present study implemented a meta-analysis to investigate whether bilingualism produces a lexical deficit. We formulated a series of alternative hypotheses that differed in their predictions as to the consistency of the lexical deficit across studies and populations. The results show that across the analyzed studies, the lexical deficit was present or more salient only under certain conditions. Crucially, these conditions were related to language learning history. Importantly, no evidence of a lexical deficit was found for simultaneous bilinguals who acquired both languages from birth. In bilinguals who acquired one language from birth and another language after that (sequential bilinguals) we found evidence of a lexical deficit under specific conditions. There was no evidence of lexical deficit in sequential bilinguals when tested in L1. A lexical deficit was present only when tested in L2. The extent of this deficit depended on age of L2 acquisition, such that the deficit was smaller in early (vs. late) learners. In addition, the lexical deficit appeared more strongly in picture-naming tasks than in fluency tasks. This result was expected because, compared to picture naming, verbal fluency (especially letter fluency) relies to a greater extent on processes other than lexical access (e.g., executive functions related to developing and maintaining retrieval strategies, see Baldo & Shimamura, 1998; Shao et al., 2014).

Overall, our findings reject the notion that bilingualism automatically brings about deviations in lexical behaviour. This interpretation is corroborated by the result that a lexical deficit was only present when sequential bilinguals were tested in their L2, not in their L1. These findings point to the possibility that the lexical deficit may be an artefact of L2 acquisition, not bilingualism. In several study designs of published research, the performance of bilingual participants tested in their L2 is compared to that of monolinguals tested in their L1. In such cases, the confounding of bilingualism versus monolingualism status with L2 versus L1 status undermines any conclusions about bilingualism being the driving factor behind verbal behaviour. Studies on the mastery of other domains of language (e.g., syntax and phonology) show that when L2 status and bilingualism are disentangled, L2 status comes out as a primary determinant of verbal behaviour (Bylund et al., 2021; Norrman & Bylund, 2016; Veríssimo et al., 2018). The current results on bilingualism type and testing language are consistent with this picture.

It would seem, then, that the findings corroborate Hypothesis 2, which predicted that language learning history would be a determinant of bilingual lexical behaviour. By extension, these results lend support to the timing-of-exposure accounts of language development, which attribute a key role to acquisition onset for the attainment of nativelike behaviour (e.g., Choi et al., 2018; Hernandez & Li, 2007; Li, 2009; Werker & Hensch, 2015). Further support for the timing-of-exposure accounts was obtained through a moderator analysis of age L2 of acquisition, which showed that individuals who started their L2 acquisition during the first decade of life were less likely to exhibit a lexical deficit than those who started past this point.

The moderator of language exposure had no measurable impact on the lexical deficit. Admittedly, it is possible that any effects of exposure were obscured by the nature of the data: because the data on this moderator were recorded at group level, the exposure values entered do not capture the actual distribution of this variable. Be that as it may, the current findings are still informative for Hypothesis 1 and, relatedly, for the amount-of-exposure accounts. The absence of a lexical deficit in simultaneous bilingualism and in L1 testing conditions shows that reduced exposure to each language, often regarded as an integral part of the bilingual experience, may not consistently incur linguistic costs.

The finding that language learning history plays an important role for lexical behaviour suggests that evidence of a lexical deficit in individuals who speak multiple languages is by and large uninformative unless learning history has been taken into account. In the absence of language learning history distinctions, it is impossible to know what underlies the observed lexical behaviour. For this reason, it is noteworthy that nearly half of the studies included in the meta-analysis either did not record learning history information or recorded the information but did not factor it into the study design. It could of course be argued that such information may have little bearing on the particular research question pursued by a given study. However, the picture that is emerging from several recent meta-analyses (besides the one at hand) is that factors of language learning history may exert a significant effect on verbal and cognitive behaviour, and neural representation (Donnelly et al., 2019; Garcia et al., 2021; Kuzmina et al., 2019; Sulpizio et al., 2020). For this reason, future studies may wish to record more details on learning history (which can easily be done with readily available background questionnaires, e.g., Anderson et al., 2017; Gullberg & Indefrey, 2003; Li et al., 2014; Marian et al., 2007) so as to anchor their findings on the characteristics of the study participants. Additionally, including this information will allow for powerful meta-analyses in the future.

Methodological considerations

There are some methodological limitations that must be considered when interpreting the outcome of the current study. Firstly, we did not include unpublished data, which would have allowed for direct assessment of the extent of publication bias in the field. Instead, we relied on statistical assessments of bias by extrapolating the association between precision (determined by sample size and measurement error) and effect size. Studies have shown that methods to assess for publication bias, including the PET-PEESE method, are prone to errors (Carter et al., 2019), and our adjusted estimates should therefore be interpreted with some caution. Secondly, we limited the included studies to those published in English, which means that our findings are not based on a complete review of existing data. Thirdly, there was considerable heterogeneity in the distribution of included effect sizes. It is possible that the current study overlooked some important moderators.

With respect to publication bias, we found that, for studies with low precision, there was some evidence of an absence of weak or negative effect sizes that would be expected as a result of sampling error. This could mean that small studies challenging the hypothesized effect were less likely to be published than small studies supporting the hypothesized effect. The association between precision and outcome remained even after removing outliers. However, the slope was not very steep. Therefore, effect sizes largely remained positive, albeit smaller, after adjusting for the observed bias.

As a recommendation for future studies, we suggest a stronger emphasis on publishing preregistered reports, because doing so decreases publication bias against null findings. We also recommend employing more open science practices, like sharing data files and analysis scripts to facilitate reanalyses and increasing data availability for future meta-analyses (see Dal Ben et al., 2022). We also call for more stringent reporting of data, including sufficient data to calculate effect sizes and confidence intervals (i.e., sample size, means, and standard deviations for each group and outcome), as well as sufficient sample descriptives. As mentioned above, for approximately half of the included samples, we could not find sufficient information to categorize bilinguals as simultaneous or sequential.

Conclusion

The current findings, based on a meta-analysis of 130 individual studies with 478 comparisons of monolinguals and bilinguals, show that bilingualism does not automatically incur a linguistic cost in the form of a lexical deficit. Rather, such a deficit may be a second language phenomenon. Because L2 speakers are often bilingual (i.e., to the extent that they continue using their mother tongue alongside the new language), the lexical deficit may simply *correlate* with the presence of bilingualism without there being any *causal* relationship between them. In the recent past, a case has been made for the importance of distinguishing between different types of bilingual experiences in order to understand their consequences on the human mind (e.g., Bialystok, 2017). However, until now, this argument primarily concerned the effects of bilingualism on cognitive functioning. As our study shows, it is of equal importance to take account of language learning history in order to understand the effects of bilingualism on linguistic functioning. This insight is hardly unexpected, because linguistic behaviour in the present is after all a product of language experiences in the past.

In view of both the current results and the recent findings on the cognitive advantage, we venture to say that our understanding of bilingualism is presently undergoing radical changes. The lexical deficit and the cognitive advantage might no longer be considered signatures of the bilingual experience. What is exciting about this development are the new insights it can bring regarding the potential and limits of the human capacity for language and its relationship to cognitive functioning.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13423-022-02136-7.

Acknowledgments We are grateful to the anonymous reviewers of *PBR* for their insightful comments on an earlier version of the manuscript. We thank Ane Theimann and Sarah Cameron for assistance in article screening.

Funding Open access funding provided by Stockholm University. This research was funded by the Olle Engkvist foundation (*Olle Engkvists stiftelse*, 200-0676, E.B. and N.A.). A.M.H.O. and M.L. have been partly supported by the Research Council of Norway through its Centers of Excellence funding scheme (223265).

References

References marked with an asterisk indicate studies included in the meta-analysis.

Abrahamsson, N. (2012). Age of onset and nativelike L2 ultimate attainment of morphosyntactic and phonetic intuition. *Studies in Second Language Acquisition, 34*(02), 187–214. https://doi.org/10.1017/S0272263112000022

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*(4), 481–509.

*Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., & Green, D. W. (2013a). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex*, *49*(3), 905–911.

*Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., Cappa, S. F., & Costa, A. (2012). Bilingualism Tunes the Anterior Cingulate Cortex for Conflict Monitoring. *Cerebral Cortex*, *22*(9), 2076–2086.

*Abutalebi, J., Della Rosa, P. A., Castro Gonzaga, A. K., Keim, R., Costa, A., & Perani, D. (2013b). The role of the left putamen in multilingual language production. *Brain and Language*, *125*(3), 307–315.

*Akhavan, N., Blumenfeld, H. K., & Love, T. (2020). Auditory sentence processing in bilinguals: The role of cognitive control. *Frontiers in Psychology*, *11*, 898.

*Anderson, J. A. E., Chung-Fat-Yim, A., Bellana, B., Luk, G., & Bialystok, E. (2018). Language and cognitive control networks in bilinguals and monolinguals. *Neuropsychologia*, *117*, 352–363.

Anderson, J. A. E., Hawrylewicz, K., & Grundy, J. G. (2020). Does bilingualism protect against dementia? A meta-analysis. *Psychonomic Bulletin & Review, 27*(5), 952–965.

*Anderson, J. A. E., Saleemi, S., & Bialystok, E. (2017). Neuropsychological assessments of cognitive aging in monolingual and bilingual older adults. *Journal of Neurolinguistics*, *43*, 17–27.

*Ansaldo, A. I., Ghazi-Saidi, L., & Adrover-Roig, D. (2015). Interference control in elderly bilinguals: Appearances can be misleading. *Journal of Clinical and Experimental Neuropsychology*, *37*(5), 455–470.

*Ardila, A., Rosselli, M., Ortega, A., Lang, M., & Torres, V. L. (2019). Oral and written language abilities in young Spanish/English bilinguals. *International Journal of Bilingualism*, *23*(1), 296–312.

*Baladzhaeva, L., & Laufer, B. (2018). Is first language attrition possible without second language knowledge? *International Review of Applied Linguistics in Language Teaching*, *56*(2), 103–136.

Baldo, J. V., & Shimamura, A. P. (1998). Letter and category fluency in patients with frontal lobe lesions. *Neuropsychology, 12*(2), 259–267.

*Barbu, C.-A., Gillet, S., & Poncelet, M. (2020). Investigating the effects of language-switching frequency on attentional and executive functioning in proficient bilinguals. *Frontiers in Psychology*, *11*, 1078.

*Baus, C., Santesteban, M., Runnqvist, E., Strijkers, K., & Costa, A. (2020). Characterizing lexicalization and self-monitoring processes in bilingual speech production. *Journal of Neurolinguistics*, *56*, 100934.

*Bellegarda, M., & Macizo, P. (2021). Cognitive control and bilingualism: The bilingual advantage through the lens of dimensional overlap. *Frontiers in Psychology*, *12*, 173.

*Bennett, J., & Verney, S. P. (2019). Linguistic factors associated with phonemic fluency performance in a sample of bilingual Hispanic undergraduate students. *Applied Neuropsychology: Adult*, *26*(4), 297–310.

*Berroir, P., Ghazi-Saidi, L., Dash, T., Adrover-Roig, D., Benali, H., & Ansaldo, A. I. (2017). Interference control at the response level: Functional networks reveal higher efficiency in the bilingual brain. *Journal of Neurolinguistics*, *43*, 4–16.

*Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, *143*(3), 233–262.

*Bialystok, E., Craik, F. I. M., & Luk, G. (2008a). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics*, *21*(6), 522–538.

Bialystok, E., Craik, F., & Luk, G. (2008b). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 859.

Bialystok, E., & Werker, J. F. (2017). Editorial: The systematic effects of bilingualism on children's development. *Developmental Science, 20*(1), e12535.

*Bice, K., & Kroll, J. F. (2021). Grammatical processing in two languages: How individual differences in language experience and cognitive abilities shape comprehension in heritage bilinguals. *Journal of Neurolinguistics*, *58*, 100963.

*Blumenfeld, H. K., Bobb, S. C., & Marian, V. (2016). The role of language proficiency, cognate status and word frequency in the assessment of Spanish–English bilinguals' verbal fluency. *International Journal of Speech-Language Pathology*, *18*(2), 190–201.

*Blumenfeld, H. K., & Marian, V. (2014). Cognitive control in bilinguals: Advantages in Stimulus–Stimulus inhibition. *Bilingualism: Language and Cognition*, *17*(3), 610–629.

*Bogulski, C. A., Bice, K., & Kroll, J. F. (2019). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition*, *22*(5), 1052–1067.

*Bogulski, C. A., Rakoczy, M., Goodman, M., & Bialystok, E. (2015). Executive control in fluent and lapsed bilinguals. *Bilingualism: Language and Cognition*, *18*(3), 561–567.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. John Wiley & Sons.

*Borodkin, K., Kenett, Y. N., Faust, M., & Mashal, N. (2016). When pumpkin is closer to onion than to squash: The structure of the second language lexicon. *Cognition*, *156*, 60–70.

Botezatu, M. R., Miller, C. A., Johnson, J., & Misra, M. (2021). Event-related potentials reveal that bilinguals are more efficient in resolving conflict than monolinguals. *NeuroReport, 32*(8), 721–726.

*Bradley, K. A. L., King, K. E., & Hernandez, A. E. (2013). Language experience differentiates prefrontal and subcortical activation of the cognitive control network in novel word learning. *NeuroImage*, *67*, 101–110.

*Broos, W. P., Duyck, W., & Hartsuiker, R. J. (2019). Monitoring speech production and comprehension: Where is the second-language delay? *Quarterly Journal of Experimental Psychology*, *72*(7), 1601–1619.

*Broos, W. P. J., Bencivenni, A., Duyck, W., & Hartsuiker, R. J. (2021). Delayed picture naming in the first and second language. *Bilingualism: Language and Cognition*, *24*(2), 389–400.

*Bylund, E., Abrahamsson, N., Hyltenstam, K., & Norrman, G. (2019). Revisiting the bilingual lexical deficit: The impact of age of acquisition. *Cognition*, *182*, 45–49.

Bylund, E., Hyltenstam, K., & Abrahamsson, N. (2021). Age of acquisition—not bilingualism—is the primary determinant of less than nativelike L2 ultimate attainment. *Bilingualism: Language and Cognition, 24*(1), 18–30.

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science, 2*(2), 115–144.

Choi, D., Black, A. K., & Werker, J. F. (2018). Cascading and multisensory influences on speech perception development. *Mind, Brain, and Education, 12*(4), 212–223.

*Clare, L., Whitaker, C. J., Martyr, A., Martin-Forbes, P. A., Bastable, A. J., Pye, K. L., Quinn, C., Thomas, E. M., Mueller Gathercole, V. C., & Hindle, J. V. (2016). Executive control in older Welsh monolinguals and bilinguals. *Journal of Cognitive Psychology*, *28*(4), 412–426.

*Claussenius-Kalman, H., Vaughn, K. A., Archila-Suerte, P., & Hernandez, A. E. (2020). Age of acquisition impacts the brain differently depending on neuroanatomical metric. *Human Brain Mapping*, *41*(2), 484–502.

*Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283.

*Costa, A., Kovacic, D., Franck, J., & Caramazza, A. (2003). On the autonomy of the grammatical gender systems of the two languages of a bilingual. *Bilingualism: Language and Cognition*, *6*(3), 181–200.

*Degani, T., & Tokowicz, N. (2013). Cross-language influences: Translation status affects intraword sense relatedness. *Memory & Cognition*, *41*(7), 1046–1064.

Dal Ben, R., Brouillard, M., Gonzalez-Barrero, A., Killam, H., Kremin, L., Quirk, E., Sander-Montant, E., Schott, E., Tsui, R. K.-Y., & Byers-Heinlein, K. (2022). *How open science can benefit bilingualism research: A lesson in six tales* (pp. 1–8). *Bilingualism*. https://doi.org/10.1017/S1366728922000256

*de Bruin, A., Della Sala, S., & Bak, T. H. (2016). The effects of language use on lexical processing in bilinguals. *Language, Cognition and Neuroscience*, *31*(8), 967–974.

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive Advantage in Bilingualism: An Example of Publication Bias? *Psychological Science, 26*(1), 99–107.

Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T., Riedel, M. C., Laird, A. R., & Gonzalez, R. (2019). No evidence for a bilingual executive function advantage in the ABCD study. *Nature Human Behaviour, 3*(7), 692–701.

Donnelly, S., Brooks, P. J., & Homer, B. D. (2019). Is there a bilingual advantage on interference-control tasks? A multiverse meta-analysis of global reaction time and interference cost. *Psychonomic Bulletin & Review, 26*(4), 1122–1147.

*Egan, C., Oppenheim, G. M., Saville, C., Moll, K., & Jones, M. W. (2019). Bilinguals apply language-specific grain sizes during sentence reading. *Cognition*, *193*, 104018.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629–634.

*Emmorey, K., Petrich, J. A. F., & Gollan, T. H. (2013). Bimodal bilingualism and the frequency-lag hypothesis. *The Journal of Deaf Studies and Deaf Education*, *18*(1), 1–11.

*Estanga, A., Ecay-Torres, M., Ibañez, A., Izagirre, A., Villanua, J., Garcia-Sebastian, M., Iglesias Gaspar, M. T., Otaegui-Arrazola, A., Iriondo, A., Clerigue, M., & Martinez-Lage, P. (2017). Beneficial effect of bilingualism on Alzheimer's disease CSF biomarkers and cognition. *Neurobiology of Aging*, *50*, 144–151.

*Fernandez, M., Tartar, J. L., Padron, D., & Acosta, J. (2013). Neurophysiological marker of inhibition distinguishes language groups on a non-linguistic executive function test. *Brain and Cognition*, *83*(3), 330–336.

*Filippi, R., Ceccolini, A., & Bright, P. (2022). Trajectories of verbal fluency and executive functions in multilingual and monolingual children and adults: A cross-sectional study. *Quarterly Journal of Experimental Psychology*, *75*(1), 130–147.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*(1), 78–104.

*Francis, W. S., Taylor, R. S., Gutiérrez, M., Liaño, M. K., Manzanera, D. G., & Penalver, R. M. (2018). The effects of bilingual language proficiency on recall accuracy and semantic clustering in free recall output: Evidence for shared semantic associations across languages. *Memory*, *26*(10), 1364–1378.

*Friesen, D. C., Luo, L., Luk, G., & Bialystok, E. (2015). Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience*, *30*(3), 238–250.

Garcia, O., Faghihi, N., Raola, A. R., & Vaid, J. (2021). Factors influencing bilinguals' speed and accuracy of number judgments across languages: A meta-analytic review. *Journal of Memory and Language, 118*, 104211.

*Giezen, M. R., & Emmorey, K. (2017). Evidence for a bimodal bilingual disadvantage in letter fluency. *Bilingualism: Language and Cognition*, *20*(1), 42–48.

*Gollan, T. H., & Brown, A. S. (2006). From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology: General*, *135*(3), 462.

*Gollan, T. H., Montoya, R. I., & Bonanni, M. P. (2005a). Proper Names Get Stuck on Bilingual and Monolingual Speakers' Tip of the Tongue Equally Often. *Neuropsychology*, *19*(3), 278.

*Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814.

*Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005b). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, *33*(7), 1220–1234.

*Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish–English bilinguals. *Neuropsychology*, *16*(4), 562.

*Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, *140*(2), 186.

*Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *15*(3), 594–615.

Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research, 29*(3), 311–343.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*(2), 131–149.

Gullberg, M., & Indefrey, P. (2003). *Gullberg Indefrey Language Background Questionnaire (English), 2003*. https://doi.org/10.13140/RG.2.2.21793.63843

Hernandez, A. E., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin, 133*(4), 638–650.

Hernandez, A., Li, P., & Macwhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences, 9*(5), 220–225.

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review, 18*(4), 625–658.

*Hirsh, K. W., Morrison, C. M., Gaset, S., & Carnicer, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition*, *6*(2), 117–128.

Hyltenstam, K., Bylund, E., Abrahamsson, N., & Park, H.-S. (2009). Dominant-language replacement: The case of international adoptees. *Bilingualism: Language and Cognition, 12*(2), 121–140.

*Ivanova, I., Salmon, D. P., & Gollan, T. H. (2013). The multilingual naming test in Alzheimer's disease: Clues to the origin of naming impairments. *Journal of the International Neuropsychological Society, 19*(3), 272–283.

*Johns, B. T., Sheppard, C. L., Jones, M. N., & Taler, V. (2016). The role of semantic diversity in word recognition across aging and bilingualism. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00703

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60–99.

*Kalia, V., Wilbourn, M. P., & Ghio, K. (2014). Better early or late? Examining the influence of age of exposure and language proficiency on executive function in early and late bilinguals. *Journal of Cognitive Psychology*, *26*(7), 699–713.

*Kan, P. F., & Sadagopan, N. (2014). Novel word retention in bilingual and monolingual speakers. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01024

*Kasparian, K., & Steinhauer, K. (2016). Confusing similar words: ERP correlates of lexical-semantic processing in first language attrition and late second language acquisition. *Neuropsychologia*, *93*, 200–217.

*Kasparian, K., & Steinhauer, K. (2017). When the second language takes the lead: Neurocognitive processing changes in the first language of adult attriters. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00389

*Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2011). The relationship between vocabulary and short-term memory measures in monolingual and bilingual speakers: *International Journal of Bilingualism*. https://doi.org/10.1177/1367006911403201

*Kaushanskaya, M., & Marian, V. (2009a). Bilingualism reduces native-language interference during novel-word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 829.

*Kaushanskaya, M., & Marian, V. (2009b). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705–710.

*Kharkhurin, A. V. (2008). The effect of linguistic proficiency, age of second language acquisition, and length of exposure to a new cultural environment on bilinguals' divergent thinking. *Bilingualism: Language and Cognition*, *11*(2), 225–243.

Kharkhurin, A. V. (2010a). Bilingual verbal and nonverbal creative behavior: *International Journal of Bilingualism*. https://doi.org/10.1177/1367006910363060

*Kharkhurin, A. V. (2010b). Sociocultural differences in the relationship between bilingualism and creative potential. *Journal of Cross-Cultural Psychology*. 10.1177/0022022110361777

*Kharkhurin, A. V. (2009). The role of bilingualism in creative performance on divergent thinking and invented alien creatures tests. *The Journal of Creative Behavior*, *43*(1), 59–71.

*Kharkhurin, A. V. (2017). Language mediated concept activation in bilingual memory facilitates cognitive flexibility. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01067

*Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. *Frontiers in Behavioral Neuroscience*, *8*. https://doi.org/10.3389/fnbeh.2014.00250

*Kovelman, I., Shalinsky, M. H., White, K. S., Schmitt, S. N., Berens, M. S., Paymer, N., & Petitto, L.-A. (2009). Dual language use in sign-speech bimodal bilinguals: FNIRS brain-imaging evidence. *Brain and Language, 109*(2), 112–123.

*Kreiner, H., & Degani, T. (2015). Tip-of-the-tongue in a second language: The effects of brief first-language exposure and long-term use. *Cognition*, *137*, 106–114.

Kuzmina, E., Goral, M., Norvik, M., & Weekes, B. S. (2019). What influences language impairment in bilingual aphasia? A meta-analytic review. *Frontiers in Psychology, 10*, 445.

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin, 144*(4), 394–425.

*Li, C., Goldrick, M., & Gollan, T. H. (2017). Bilinguals' twisted tongues: Frequency lag or interference? *Memory & Cognition*, *45*(4), 600–610.

*Li, L., Abutalebi, J., Zou, L., Yan, X., Liu, L., Feng, X., Wang, R., Guo, T., & Ding, G. (2015). Bilingualism alters brain functional connectivity between "control" regions and "language" regions: Evidence from bimodal bilinguals. *Neuropsychologia*, *71*, 236–247.

Li, P. (2009). Lexical Organization and Competition in First and Second Languages: Computational and Neural Mechanisms. *Cognitive Science, 33*(4), 629–664.

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition, 17*(3), 673–680.

*Li, Y., Yang, J., Suzanne Scherf, K., & Li, P. (2013). Two faces, two languages: An fMRI study of bilingual picture naming. *Brain and Language*, *127*(3), 452–462.

*Lin, Y.-C., & Lin, P.-Y. (2016). Mouse tracking traces the "Camrbidge Unievrsity" effects in monolingual and bilingual minds. *Acta Psychologica*, *167*, 52–62.

*Ljungberg, J. K., Elbe, P., & Sörman, D. E. (2020). The bilingual effects of linguistic distances on episodic memory and verbal fluency. *Scandinavian Journal of Psychology*, *61*(2), 195–203.

*Ljungberg, J. K., Hansson, P., Andrés, P., Josefsson, M., & Nilsson, L.-G. (2013). A Longitudinal Study of Memory Advantages in Bilinguals. *PLOS ONE*, *8*(9), e73029.

*Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, *114*(1), 29–41.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940–967.

*Massa, E., Köpke, B., & El Yagoubi, R. (2020). Age-related effect on language control and executive control in bilingual and monolingual speakers: Behavioral and electrophysiological evidence. *Neuropsychologia*, *138*, 107336.
*Milman, L. H., Faroqi-Shah, Y., Corcoran, C. D., & Damele, D. M. (2018). Interpreting Mini-Mental State Examination performance in highly proficient bilingual Spanish–English and Asian Indian–English Speakers: Demographic adjustments, item analyses, and supplemental measures. *Journal of Speech, Language, and Hearing Research*, *61*(4), 847–856.
*Misdraji-Hammond, E., Lim, N. K., Fernandez, M., & Burke, M. E. (2015). Object familiarity and acculturation do not explain performance difference between Spanish–English bilinguals and English monolinguals on the Boston Naming Test. *Archives of Clinical Neuropsychology*, *30*(1), 59–67.
*Mizrahi, R., Wixted, J. T., & Gollan, T. H. (2021). Order effects in bilingual recognition memory partially confirm predictions of the frequency-lag hypothesis. *Memory*, *29*(4), 444–455.
*Mor, B., Yitzhaki-Amsalem, S., & Prior, A. (2015). The joint effect of bilingualism and ADHD on executive functions. *Journal of Attention Disorders*, *19*(6), 527–541.
*Morrison, C., & Taler, V. (2020). ERP measures of the effects of age and bilingualism on working memory performance. *Neuropsychologia*, *143*, 107468.
*Navarro-Torres, C. A., Garcia, D. L., Chidambaram, V., & Kroll, J. F. (2019). Cognitive control facilitates attentional disengagement during second language comprehension. *Brain Sciences*, *9*(5), 95.
Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., & Owen, A. M. (2020). Bilingualism affords no general cognitive advantages: A population study of executive function in 11,000 people. *Psychological Science, 31*(5), 548–567.
Norrman, G., & Bylund, E. (2016). The irreversibility of sensitive period effects in language development: Evidence from second language acquisition in international adoptees. *Developmental Science, 19*(3), 513–520.
*Olabarrieta-Landa, L., Benito-Sánchez, I., Alegret, M., Gailhajanet, A., Landa Torre, E., López-Mugartza, J. C., & Arango-Lasprilla, J. C. (2019). Letter verbal fluency in Spanish-, Basque-, and Catalan-speaking individuals: Does the selection of the letters influence the outcome? *Journal of Speech, Language, and Hearing Research*, *62*(7), 2400–2410.
*Olsen, R. K., Pangelinan, M. M., Bogulski, C., Chakravarty, M. M., Luk, G., Grady, C. L., & Bialystok, E. (2015). The effect of lifelong bilingualism on regional grey and white matter volume. *Brain Research*, *1612*, 128–139.
*Oschwald, J., Schättin, A., Bastian, C. C. von, & Souza, A. S. (2018). Bidialectalism and bilingualism: Exploring the role of language similarity as a link between linguistic ability and executive control. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01997
Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232–258.
*Paap, K. R., & Liu, Y. (2014). Conflict resolution in sentence processing is the same for bilinguals and monolinguals: The role of confirmation bias in testing for bilingual advantages. *Journal of Neurolinguistics, 27*(1), 50–74.
*Paap, K. R., Myuz, H. A., Anders, R. T., Bockelman, M. F., Mikulinsky, R., & Sawi, O. M. (2017). No compelling evidence for a bilingual advantage in switching or that frequent language switching reduces switch cost. *Journal of Cognitive Psychology*, *29*(2), 89–112.
Palmer, T. M., Sutton, A. J., Peters, J. L., & Moreno, S. G. (2008). Contour-enhanced funnel plots for meta-analysis. *The Stata Journal, 8*(2), 242–254.
*Palomar-García, M.-Á., Bueichekú, E., Ávila, C., Sanjuán, A., Strijkers, K., Ventura-Campos, N., & Costa, A. (2015). Do bilinguals show neural differences with monolinguals when processing their native language? *Brain and Language*, *142*, 36–44.
*Paplikar, A., Alladi, S., Varghese, F., Mekala, S., Arshad, F., Sharma, M., Saroja, A. O., Divyaraj, G., Dutt, A., Ellajosyula, R., Ghosh, A., Iyer, G. K., Sunitha, J., Kandukuri, R., Kaul, S., Khan, A. B., Mathew, R., Menon, R., Nandi, R., … ICMR-NCTB Consortium. (2021). Bilingualism and its implications for neuropsychological evaluation. *Archives of Clinical Neuropsychology*, *36*(8), 1511–1522.
Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. John Benjamins.
*Patra, A., Bose, A., & Marinis, T. (2020). Performance difference in verbal fluency in bilingual and monolingual speakers. *Bilingualism: Language and Cognition*, *23*(1), 204–218.
*Pelham, S. D., & Abrams, L. (2013). Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 313.
*Peñaloza, C., Grasemann, U., Dekhtyar, M., Miikkulainen, R., & Kiran, S. (2019). BiLex: A computational approach to the effects of age of acquisition and language exposure on bilingual lexical access. *Brain and Language*, *195*, 104643.
*Peristeri, E., Tsimpli, I. M., Sorace, A., & Tsapkini, K. (2018). Language interference and inhibition in early and late successive bilingualism. *Bilingualism: Language and Cognition*, *21*(5), 1009–1034.
Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology, 61*(10), 991–996.
*Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, *22*(3), 415–422.
*Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society*, *17*(4), 682–691.
*Pyers, J. E., Gollan, T. H., & Emmorey, K. (2009). Bimodal bilinguals reveal the source of tip-of-the-tongue states. *Cognition*, *112*(2), 323–329.
R Core Team. (2017). R—A language and environment for statistical computing. In *R Foundation for Statistical Computing*. Vienna Austria: R Version 3.4.3.
Rivera-Gaxiola, M., Klarman, L., Garcia-Sierra, A., & Kuhl, P. K. (2005). Neural patterns to speech and vocabulary growth in American infants. *NeuroReport, 16*(5), 495–498.
*Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology*, *16*(4-6), 635–645.
*Rodriguez-Fornells, A., Lugt, A. van der Rotte, M., Britti, B., Heinze, H.-J., & Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: Brain potential and functional imaging evidence. *Journal of Cognitive Neuroscience*, *17*(3), 422–433.
*Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solí, F. (2000). Verbal fluency and repetition skills in healthy older Spanish–English bilinguals. *Applied Neuropsychology*, *7*(1), 17–24.
*Rosselli, M., Ardila, A., Lalwani, L. N., & Vélez-Uribe, I. (2016). The effect of language proficiency on executive functions in balanced and unbalanced Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *19*(3), 489–503.
*Rosselli, M., Ardila, A., Salvatierra, J., Marquez, M., Luis, M., & Weekes, V. A. (2002). A cross-linguistic comparison of verbal

fluency tests. *International Journal of Neuroscience*, *112*(6), 759–776.

*Runnqvist, E., Gollan, T. H., Costa, A., & Ferreira, V. S. (2013). A disadvantage in bilingual sentence production modulated by syntactic frequency and similarity across languages. *Cognition*, *129*(2), 256–263.

*Ryskin, R. A., Brown-Schmidt, S., Canseco-Gonzalez, E., Yiu, L. K., & Nguyen, E. T. (2014). Visuospatial perspective-taking in conversation and the role of bilingual experience. *Journal of Memory and Language*, *74*, 46–76.

*Sadat, J., Martin, C. D., Alario, F. X., & Costa, A. (2012). Characterizing the bilingual disadvantage in noun phrase production. *Journal of Psycholinguistic Research*, *41*(3), 159–179.

*Sadat, J., Martin, C. D., Magnuson, J. S., Alario, F.-X., & Costa, A. (2016). Breaking down the bilingual cost in speech production. *Cognitive Science*, *40*(8), 1911–1940.

*Salvatierra, J. L., & Rosselli, M. (2010). The effect of bilingualism and age on inhibitory control: *International Journal of Bilingualism*. https://doi.org/10.1177/1367006910371021

*Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, *13*(2), 231–252.

*Sasisekaran, J., & Weisberg, S. (2013). The effects of cognitive: Linguistic variables and language experience on behavioural and kinematic performances in nonword learning. *Journal of Psycholinguistic Research*, *42*(2), 175–190.

*Savoie, J., Root, K., Villers, J., Goldsmith, K. M., & Short, M. (2019). Boston Naming Test performance in French-speaking Acadians. *Aphasiology*, *33*(5), 561–578. 10.1080/02687038.2018.1490387

*Schmid, M. S. (2014). The debate on maturational constraints in bilingual development: A perspective from first-language attrition. *Language Acquisition*, *21*(4), 386–410.

*Schmidtke, J. (2014). Second language experience modulates word retrieval effort in bilinguals: Evidence from pupillometry. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00137

*Schmidtke, J. (2016). The bilingual disadvantage in speech understanding in noise is likely a frequency effect related to reduced language exposure. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00678

*Seçer, I. (2016). Skills of cognitive flexibility in monolingual and bilingual younger adults. *The Journal of General Psychology*, *143*(3), 172–184.

*Segal, D., & Gollan, T. H. (2018). What's left for balanced bilinguals? Language proficiency and item familiarity affect left-hemisphere specialization in metaphor processing. *Neuropsychology*, *32*(7), 866.

*Sehyr, Z. S., Giezen, M. R., & Emmorey, K. (2018). Comparing semantic fluency in American Sign Language and English. *The Journal of Deaf Studies and Deaf Education*, *23*(4), 399–407.

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00772

*Sheppard, C., Kousaie, S., Monetta, L., & Taler, V. (2016). Performance on the Boston Naming Test in bilinguals. *Journal of the International Neuropsychological Society*, *22*(3), 350–363.

Shook, A., & Marian, V. (2012). Bimodal bilinguals co-activate both languages during spoken comprehension. *Cognition, 124*(3), 314–324.

*Soltani, M., Moradi, N., Rezaei, H., Hosseini, M., & Jasemi, E. (2019). Comparison of verbal fluency in monolingual and bilingual elderly in Iran. *Applied Neuropsychology: Adult*, *0*(0), 1–8. https://doi.org/10.1080/23279095.2019.1594234

Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin & Review, 24*(4), 1077–1096.

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5*(1), 60–78.

*Stasenko, A., & Gollan, T. H. (2019). Tip of the tongue after any language: Reintroducing the notion of blocked retrieval. *Cognition*, *193*, 104027.

*Stasenko, A., Matt, G. E., & Gollan, T. H. (2017). A relative bilingual advantage in switching with preparation: Nuanced explorations of the proposed association between bilingualism and task switching. *Journal of Experimental Psychology: General*, *146*(11), 1527.

*Sullivan, M. D., Poarch, G. J., & Bialystok, E. (2018). Why is lexical retrieval slower for bilinguals? Evidence from picture naming. *Bilingualism: Language and Cognition*, *21*(3), 479–488.

Sulpizio, S., Del Maschio, N., Fedeli, D., & Abutalebi, J. (2020). Bilingual language processing: A meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 108*, 834–853.

*Sundaray, S., Marinis, T., & Bose, A. (2018). Comprehending nonliteral language: Effects of aging and bilingualism. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02230

*Taler, V., Johns, B. T., Young, K., Sheppard, C., & Jones, M. N. (2013). A computational analysis of semantic structure in bilingual verbal fluency performance. *Journal of Memory and Language*, *69*(4), 607–618.

*Taler, V., López Zunini, R., & Kousaie, S. (2016). Effects of Semantic Richness on Lexical Processing in Monolinguals and Bilinguals. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00382

*Tao, L., Taft, M., & Gollan, T. H. (2015). The Bilingual switching advantage: Sometimes related to bilingual proficiency, sometimes not. *Journal of the International Neuropsychological Society*, *21*(7), 531–544.

*Vega-Mendoza, M., West, H., Sorace, A., & Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. *Cognition*, *137*, 40–46.

Veríssimo, J., Heyer, V., Jacob, G., & Clahsen, H. (2018). Selective effects of age of acquisition on morphological priming: Evidence for a sensitive period. *Language Acquisition, 25*(3), 315–326.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3). 10.18637/jss.v036.i03

Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology, 66*(1), 173–196.

*Woumans, E., Ceuleers, E., Van der Linden, L., Szmalec, A., & Duyck, W. (2015). Verbal and nonverbal cognitive control in bilinguals and interpreters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(5), 1579. https://doi.org/10.1037/xlm0000107

*Woumans, E., Van Herck, S., & Struys, E. (2019). Shifting gear in the study of the bilingual advantage: Language Switching Examined as a Possible Moderator. *Behavioral Sciences*, *9*(8), 86.

Wu, Y. J., & Thierry, G. (2011). *Event-related brain potential investigation of preparation for speech production in late bilinguals. Frontier in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00114

*Yoo, J., & Kaushanskaya, M. (2012). Phonological memory in bilinguals and monolinguals. *Memory & Cognition*, *40*(8), 1314–1330.

*Zirnstein, M., Hell, J. G. van, & Kroll, J. F. (2018). Cognitive control ability mediates prediction costs in monolinguals and bilinguals. *Cognition*, *176*, 87–106.

*Zirnstein, M., Hell, J. G. van, & Kroll, J. F. (2019). Cognitive control and language ability contribute to online reading comprehension: Implications for older adult bilinguals. *International Journal of Bilingualism*, *23*(5), 971–985.
*Zou, L., Abutalebi, J., Zinszer, B., Yan, X., Shu, H., Peng, D., & Ding, G. (2012a). Second language experience modulates functional brain network for the native language production in bimodal bilinguals. *NeuroImage*, *62*(3), 1367–1375.
*Zou, L., Ding, G., Abutalebi, J., Shu, H., & Peng, D. (2012b). Structural plasticity of the left caudate in bimodal bilinguals. *Cortex*, *48*(9), 1197–1206.

OPEN PRACTICES STATEMENT

The R script for the reported analyses, output files, and the data file can be accessed at the project's domain at the Open Science Framework (https://osf.io/8tepa/).