Studies in Second Language Acquisition (2026), **48**, 784–807
doi:10.1017/S0272263126101806

RESEARCH ARTICLE

Anything you can say, I can say faster: The impact of item frequency, frequency of language use, and inhibitory control on monolingual and bilingual picture naming speed

Monika Schmid[1], Beatriz Martin Gascon[2], Concepción Soto[3] and Ehsan Solaimani[1]

[1]Language and Linguistic Science, University of York, York, UK; [2]Lengua Española y Teoría Literaria, Universidad Complutense de Madrid, Madrid, Spain and [3]Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, USA
Corresponding author: Monika Schmid; Email: monika.schmid@york.ac.uk

(Received 31 May 2024 UTC; Revised 16 April 2026 UTC; Accepted 17 May 2026 UTC)

Abstract

Bilinguals tend to have slower response times in picture naming than monolinguals in both their first and second languages. Naming speed interacts with item frequency, suggesting that the bilingual disadvantage may be due to less frequent exposure to each language. We measure response speed in a Picture Naming and a Picture-Word-Interference paradigm among 98 bilinguals and 83 monolinguals, looking for interactions between word frequency and frequency of use of both L1 and L2. We confirm previous findings that monolinguals outperform bilinguals, and that the bilingual disadvantage is stronger for low- than for high-frequency items. While immersed bilinguals seem to have an advantage in L2 naming over non-immersed speakers, self-reported use of either language does not affect naming in either language, and the L1 is not affected by the immersed vs. non-immersed distinction. We conclude that frequency remains an elusive factor in bilingual development, and particularly in L1 attrition.

Keywords: bilingual development; frequency effect; language attrition; language use frequency; lexical access; picture naming

Introduction

There is a considerable body of research suggesting that lexical access is more effortful for second language (L2) than first language (L1) speakers, and that this additional effort manifests itself in phenomena such as higher disfluency rates in free speech (de Jong et al., 2015), more tip-of-the-tongue states (Gollan & Acenas, 2004) and longer latencies in word naming (Botezatu, Misra, & Kroll, 2022) and picture naming (e.g., Gollan et al., 2005) in L2 use than in comparable data from L1 speakers. Similar findings

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

have been reported in the L1 of (late) bilinguals as compared to monolinguals (see, for example, Bergmann, Sprenger, & Schmid, 2015 and Schmid & Beers Fägersten, 2010, for fluency in free speech; Baus, Costa, & Carreiras, 2013, Köpke & Genevska-Hanke, 2018 and Schmid & Yılmaz, 2021, for latencies and accuracy in picture naming; and Baus, Costa, & Carreiras, 2013, Ivanova & Costa, 2008; Lehtinen, Kautto, & Renvall, 2024, Linck, Kroll, & Sunderman, 2009 and Opitz, 2010 for productivity in lexical fluency tasks; for a comprehensive overview of lexical access in first language attrition see Jarvis, 2019). Lexical access in all of a bilingual's languages further appears to be affected by relative levels of proficiency in both languages, as well as length and frequency of exposure (Baus, Costa, & Carreiras, 2013; Kasparian & Steinhauer, 2016; Linck, Kroll, & Sunderman, 2009).

Response latencies for both monolinguals and bilinguals have also consistently been shown to be affected by word frequency. For bilinguals, this factor exists in a further complex relationship with frequency of use and exposure to each language: a speaker who uses one of their languages only rarely may encounter even high-frequency words of that language less often than a prolific bilingual speaker or a monolingual encounters low-frequency ones. General effects of bilingualism in lexical access, which arise from the need to control two or more linguistic systems (Green & Abutalebi, 2013), may therefore intersect with the time afforded to the individual speaker to expose themselves to each respective language ("weaker links" hypothesis, Gollan et al., 2005, 2008). In particular, availability of L1 items appears to be reduced mostly for low-frequency non-cognate items, and this reduction can be observed even after short immersion periods of a few months (Baus, Costa, & Carreiras, 2013).

We thus hypothesize that, *ceteris paribus*, the effects of item frequency on bilingual word retrieval speed in both L1 and L2 will be modulated by the amount and context of exposure that individuals have with each of the languages in daily life, and that including bilingual experience as a predictor on the participant level can allow us to refine our understanding of how bilingualism affects lexical accessibility. While this is commonly recognized in the research design of studies of language attrition, which routinely include measures of frequency of use, it seems to not have been fully explored in investigations of lexical access among L2 speakers (see below). We therefore take an integrated approach, assessing speed of lexical access as well as language proficiency and frequency of language exposure/use for both of our bilingual participants' languages, comparing them against monolingual reference groups, to assess how general properties of linguistic items interface with individual experience-related factors.

The frequency effect in lexical access

The observation that high-frequency lexical items are accessed faster by both monolinguals and bilinguals across a range of experimental paradigms is one of the least controversial findings in the field – Gollan et al. (2005, p. 1222) succinctly conclude that "words that are produced more often are easier to produce." The effect of frequency on naming latency was first observed by Oldfield and Wingfield (1965) and has been reliably replicated ever since (e.g. Alario, Costa & Caramazza, 2002; Strijkers, Costa & Thierry, 2010).

Theoretical underpinnings for the frequency effect on the individual level are provided in psycho-/neurolinguistic models of activation, which stipulate that the accessibility of items depends, among other factors, on the frequency and recency of previous instances of activation (e.g., Paradis, 2004). This is reflected in the notion of entrenchment – the idea that (linguistic) knowledge is not static but undergoes lifelong

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

processes of reorganization and refreshment, dependent on experience and exposure (e.g., H.-J. Schmid, 2017). These models all predict that strength of representation of linguistic items, and thus the ease with which they may be activated, increases with each instance of repetition/rehearsal, while disuse will cause decay, resulting in longer activation times (e.g., Langacker, 1987; Paradis, 2007).

Ease of lexical access, however, is not only constrained by how frequently a speaker has encountered an item but also by the necessity to inhibit competitors, that is, to manage crosslinguistic interference. This factor affects monolinguals and bilinguals in a fundamentally different way, as bilinguals have to restrict the lexical access process to the target language and experience interference from competing items from the non-target one, while monolinguals only have to inhibit competitors from the target language (e.g., synonyms and homophones) (e.g., Green, 1998; Green & Abutalebi, 2013; Shao, Meyer & Roelofs, 2013). This suggests that, while lexical access should be affected quantitatively by factors related to frequency, the impact of interference should differ qualitatively between monolinguals and bilinguals. A distinction has to be made here between *nonselective* inhibition, in which *any* unwanted response is suppressed (measured, e.g., in Stop-Go tasks), and *selective inhibition*, which relates to particular distractors. This type of inhibition manifests itself in effects such as cognate facilitation (e.g., Dijkstra et al., 2010) and the Stroop-like tasks, such as the Picture Word Interference paradigm (Shao, Roelofs, Martin, & Meyer, 2015; Starreveld & La Heij, 2017). In the latter type of task, bilinguals have to apply two levels of inhibition, one relating to the suppression of the distractor item, the other at the global level of the language (Neumann, 2018); while monolinguals only experience the former.

While item frequency has been shown to affect naming latency across different types of populations, from healthy monolingual adults through early and late bilinguals in both their L1 and their L2 to participants with developmental or acquired disorders, there are some differences in how the effect plays out across populations. In particular, L2 speakers are not only slower than monolinguals overall in naming both high frequency (HF) and low frequency (LF) items, but the slope for item frequency vs. naming latency is steeper for them. In other words, the bilingual disadvantage in the speed of lexical retrieval is amplified for low-frequency items (Gollan et al., 2005, 2008; Ivanova & Costa, 2008). Such differences have been interpreted as an indirect frequency effect ascribed to the larger total number of lexical items competing for selection in the bilingual lexical repertoire, as well as the reduced total frequency of activation of each item in the bilingual as opposed to the monolingual experience. For example, the weaker-links hypothesis (Gollan et al., 2008) compares being bilingual to having a monolingual lexicon that consists entirely of low-frequency words, and in which the reduced total number of instances that lexical items in each language are called upon leads to an attenuation of the link between semantics and phonology in either system and makes activation of each item more effortful. Similarly, Ramscar et al. (2013) show that the size of the overall repertoire impacts naming speed in a complex interaction between the frequency of occurrence and the size of the vocabulary. While Ramscar et al. investigate this in the context of lifelong learning (proposing that slower naming in healthy older participants is not necessarily an indication of cognitive decline but the outcome of more experience), a similar case can be made for bilinguals.

Frequency thus exists as a variable on both the item and the speaker level: lexical items are common or rare, and this factor is relatively constant for each word across the speech community, while individual (in particular bilingual) speakers are exposed to each language more or less in their daily activities. Frequency is a complex variable for research on bilingualism, since relative item frequency presumably affects

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

monolinguals and bilinguals similarly (an item that is low-frequency in monolingual discourse is unlikely to be high-frequency for bilinguals, or vice versa), but is modulated by individual exposure to each language to different degrees. A bilingual individual interacting predominantly in Language A will have almost as many exposures to items across the frequency range as a monolingual of this language, while occurrences of even high-frequency items in Language B will be lower. A speaker who uses the languages approximately equally across their daily activities will have exposure which is reduced against the monolingual baseline in both.

Few studies of lexical access and naming latency in L2 development to date have measured bilingual exposure as an individual, scaled predictor; instead, it is usually treated as a categorical variable at the group level, differentiating bilinguals vs. monolinguals (e.g., Gollan et al., 2005, 2008) or trilinguals vs. bilinguals (Sullivan, Poarch, & Bialystok, 2018). Where bilingual populations are taken to be stratified, the different levels have also sometimes been assumed to be dichotomous, for example, in de Bruin, Della Sala, and Bak's (2016) investigation of active vs. inactive heritage bilinguals or in Linck, Kroll, and Sunderman's (2009) study of immersed vs. classroom learners. While it has been acknowledged that more research into the various mechanisms underlying bilingual processing is needed, and that frequency of use is likely among the predictors (Shook, Goldrick, Engstler, & Marian, 2015), there are few studies attempting to establish the impact of frequency of exposure as an independently measured continuous variable at the individual level on naming speed in bilinguals. One recent exception is a study by Kastenbaum et al. (2019), who extract a composite language exposure measure from a survey by means of Principal Component Analysis, and find it to be a significant predictor in a category fluency task (note that the impact of item frequency is more difficult to establish and control in such a task than in experimental paradigms where all participants produce the same items, such as picture naming).

On the other hand, there is a considerable body of work investigating frequency of exposure as a factor in the development of bilinguals' L1, the process known as L1 attrition. While the item frequency effect observed for other populations is also reliably replicated in this context (e.g., Hulsen, 2000; Yılmaz & Schmid, 2012), the impact of frequency of individual exposure on the rate and extent of attrition effects has been surprisingly elusive (see Schmid, 2019, for an overview). These findings present a puzzle for accounts of the frequency effect by means of an activation or entrenchment narrative and the "use it or lose it" scenario predicted by accounts such as Langacker (1987) or Paradis (2007).

Research questions and hypotheses

The present study seeks to contribute to an integrated picture of the impact of word frequency, frequency of exposure, proficiency, and the cognitive demands of managing two linguistic systems on naming latency and lexical access in both the L1 and the L2 of bilinguals with a broad range of experiential backgrounds, proficiency levels, and amount of exposure. We assess the impact of item-level characteristics (frequency and cognateness) and participant-level predictors (individual background factors, executive functions) on naming speed. In order to test the impact of inhibition demands, we use both a Picture Naming Task (PNT) and a Picture Word Interference Task (PWIT), in which target pictures are presented together with a distractor item in the form of a word printed across the picture, which the participant is instructed to ignore (e.g., Shao, Roelofs, Martin, & Meyer, 2015; Starreveld & La Heij, 2017;

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

Sudarshan & Baum, 2019). Picture naming is used to measure the differential effects of the background factors on the accessibility of lexical items, while testing the impact of the presence of a distractor is used to probe whether the ability to selectively inhibit a distractor item in either of the languages of the bilingual is modulated by effects of frequency of use and/or proficiency level.

We address the following research questions (see Table 1 for a summary):

RQ1 (lexical access): To what extent does naming latency in a PNT and a PWIT differ between a monolingual reference population, a bilingual population naming items in their L1, and a bilingual population naming them in their L2?

Hypothesis Ia: Bilinguals will be slower to name items in both L1 and L2 than monolinguals of these languages

Hypothesis Ib: Bilinguals will be faster to name items in their L1 than in their L2

RQ2 (bilingual language dominance): To what extent does naming latency in these tasks differ between bilingual speakers naming the items in their L1 and the same bilinguals naming the same items in their L2?

Hypothesis IIa: Differences in naming speed between bilinguals' L1 and L2 will be modulated by proficiency

Hypothesis IIb: Differences in naming speed between bilinguals' L1 and L2 will be modulated by frequency of exposure

RQ3 (inhibition): What is the cost incurred when naming an item in the presence of a (semantically and phonologically unrelated) distractor in bilinguals' L1 and L2, compared with a monolingual baseline?

Hypothesis III: Bilinguals will incur a larger cost for inhibiting the distractor word in the PWIT, as their overall inhibition effort has to extend not only to ignoring the unrelated distractor but also to suppress competitor items from the non-target language

RQ4 (individual differences): How do individual predictors relating to frequency of exposure, language proficiency, and executive function modulate lexical access, language dominance, and inhibition in bilinguals' L1 and L2 across the item frequency range?

Hypothesis IVa: Item frequency will have a stronger effect on naming latency for less frequently used languages

Table 1. Summary of RQs and Hypotheses

RQ1—lexical access
HIa: monolinguals < bilinguals
HIb: bilinguals L1 < bilinguals L2
RQ2—language dominance
HIIa: higher L2 proficiency ≥ smaller gap between L1 and L2
HIIb: frequency of exposure associated with dominance measure
RQ3—inhibition
HIII: larger PWIT cost for bilinguals
RQ4—individual differences
HIVa: interaction item frequency and frequency of use
HIVb: interaction item frequency and proficiency

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

Hypothesis IVb: Item frequency will affect participants with lower levels of proficiency more strongly

The study

Participants

Ethical approval for this investigation was obtained from the Faculty Ethics Committee of the Faculty of Social Sciences, University of Essex.[1] Informed consent was obtained before any of the tasks were attempted.

Participant recruitment was initiated through the authors' personal networks, beginning with colleagues and students in the departments with which the authors were affiliated at the time and expanding to further contacts which participants were able to provide. The experiment was completed by 181 adult participants.[2] Of these, 83 (27 males) were functionally monolingual (no or low proficiency and infrequent exposure to languages other than their L1), with 31 (13) being native speakers of German (GEMono), 30 (9) of Spanish (SPMono), and 22 (5) of English (ENMono). The remaining 98 participants were fluent in two languages and had all started to learn the second language after age 10; 46 (16) of these had English as their L1, which was the L2 for the remaining 42 (28). Twenty-three (16) were L1 speakers of German (GEL1ENL2) and 29 (12) of different varieties of Spanish (SPL1ENL2). Of the 46 English native bilinguals, 25 (10) were L2 speakers of German (ENL1GEL2) and 21 (6) of Spanish (ENL1SPL2). Education levels were uniformly high, with just over half of participants attending or having attended university. The youngest participant was 19, the oldest was 73, with an average age of 32.81 (std 11.81).[3] These characteristics are summarised in Table 2.

All bilingual participants as well as the English monolinguals resided in England at the time of testing, while monolingual Spanish and German participants were residents of Southern Spain and the Berlin region, respectively.

Prior to attending the naming experiment, participants completed an online Personal Background Questionnaire and a proficiency assessment on the MoodleX platform. This assessment consisted of a C-Test and a self-assessment (the rationale for

[1]In the revision process, following the recommendation of a reviewer, the experiment was repeated with 11 more monolingual English participants in the summer of 2025, as the original data collection had been curtailed by the pandemic (see below), leading to unbalanced participant numbers in this group. As all authors had changed affiliations in the meantime, ethical approval was again obtained from the University of York, the first author's new affiliation.

[2]The original aim had been to recruit 30 participants in each subgroup. Unfortunately, data collection was severely disrupted by the COVID-19 pandemic, leading to somewhat unbalanced numbers of participants in the different cohorts. The size of the participant pool recruited was thus smaller than we had anticipated or wished. In order to assess to what extent these limitations affect the outcomes reported here, we conducted a retrospective power analysis for the final models in Analyses 1, 2, and 3 presented below. This was done by iteratively simulating data based on the model estimates, fitting the final model for each analysis, and calculating the proportion of times the p-value of the corresponding estimate was <.05. The results indicate that despite the recruitment challenges due to COVID-19 restrictions, the study maintained good power (70%–100%) for the main effect and interactions reported in the models, although there was limited sensibility for some of the between-subject predictors and proficiency-related terms.

[3]Several reviewers have remarked on the large age range. This is a near-universal characteristic of language attrition studies, which cannot draw on populations that are as homogenous as those typically included in studies of L2 acquisition. Age was not a significant predictor in any of the final models presented below.

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

Table 2. Distribution of Gender, Age, and Education Level by Subgroup

				age		
Group	L1	L2	n (males)	mean	std	range
GEMono	German	-	31 (13)	26.97	7.46	19–48
GEL1ENL2	German	English	23 (7)	42.87	10.52	21–73
ENL1GEL2	English	German	25 (10)	42.48	13.80	23–67
SPMono	Spanish	-	30 (9)	23.30	3.94	20–39
SPL1ENL2	Spanish	English	29 (12)	33.21	8.45	20–54
ENL1SPL2	English	Spanish	21 (6)	26.10	6.87	19–48
ENMono	English	-	11 (0)	33.09	12.82	18–56

these tasks is described in detail in Schmid, 2011; the version of the tasks we used is available in the three project languages on https://languageattrition.org/resources-for-researchers/) in both of each bilinguals' languages or in the L1 of the monolinguals. The C-tests were scored by giving one point for each correctly filled-in gap[4] and then converting this score to a percentage of correctly completed gaps. The descriptive statistics in Table 3 suggest that the difficulty of the tests across the three languages may have been somewhat variable, as the Spanish version obtained higher average scores across all groups than the English or the German. We therefore calculated z-scores for the C-Test in each language and used those instead of the percentages reported here.

The Can-Do scales comprise 44 questions on participants' ability to carry out certain tasks in both of their languages. They are modeled on the Can-Do scales for proficiency levels C1 and C2 developed in the Common European Framework of Reference by the Association of Language Testers in Europe (ALTE). Participants respond to each question on a scale from 1 ("I cannot do this") to 5 ("I can do this with no effort"), with bilinguals indicating this for each of their languages. Reliability across the 44 questions was excellent (Cronbach $\alpha > .98$ for each language), so a single overall average was calculated for each of the languages of each participant. The resulting proficiency measures correlated highly in participants' L2 (Pearson's $r = .648$, $p < .001$), but not in their L1 (Pearson's $r = .029$, $p = .702$), suggesting participants were more realistic in assessing their proficiency levels in their L2 than in their L1.

Participants furthermore completed a survey on their language learning history and the use of and attitudes towards each of their languages (see Online Appendix, Table A1). This survey was presented to each participant in their L1 through the MoodleX platform and participants completed it online before attending the in-person part of the experiment. While questions were framed using the name of each language (e.g. "How often do you use English at work?") we recoded responses so that they pertained to participants' L1 or L2.

In order to explore the underlying structure of the responses collected in this survey, we followed the approach adopted by Kastenbaum et al. (2019) and Schmid & Dusseldorp (2010), who recommend reducing the number of dimensions while preserving as much as possible of the inherent variability of such a survey through a Principal Component Analysis (PCA). The following 16 variables were entered into the analysis:

- proficiency in L1 and L2 (6 variables: z-transformed C-Test in L1 and L2, Can-Do Scales in L1 and L2, Likert-scale question about L1 and L2 proficiency with 1 = very poor and 5 = very good)

[4]Omissions of diacritics were not penalized.

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

Table 3. Proficiency Measures by Subgroup

	C-test L1 (% accurate responses)		C-Test L2 (% accurate responses)		Can-Do L1 (max = 5)		Can-Do L2 (max = 5)	
Group	mean (std)	range	mean (std)	range	mean (std)	range	mean (std)	range
GEMono	86.12 (7.09)	67.44–95.35			4.82 (.19)	4.41–5		
GEL1ENL2	89.38 (6.11)	69.77–97.67	91.09 (5.16)	80–100	4.77 (.23)	4.12–5	4.66 (.35)	3.68–5
ENL1GEL2	90.5 (8.54)	72.5–100	74.33 (10.65)	53.49–93.02	4.91 (.09)	4.73–5	3.78 (.77)	2.39–4.9
SPMono	96.82 (2.53)	89.39–100			4.85 (.15)	4.59–5		
SPL1ENL2	93.1 (7.72)	63.64–100	77.76 (20.39)	17.5–100	4.83 (.21)	4.32–5	4.39 (.52)	3.1–5
ENL1SPL2	87.89 (6.19)	75–97.5	80.59 (15.01)	42.42–98.48	4.83 (.19)	4.39–5	4.06 (.62)	2.24–5
ENMono	84.89 (9.74)	65–100			4.52 (0.59)	3.34–5	.	.

- frequency of use of L1 and L2 with friends, relatives, and at work (6 Likert-scale variables, 5 = daily, 1 = very rarely or never)
- language participant feels more comfortable using (1 Likert-scale variable, 1 = strong preference for L2, 5 = strong preference for L1)
- predominant language in social networks (1 Likert-scale variable, 1 = all friends are speakers of L2, 5 = all friends are speakers of L1)
- language use with pets (2 Likert-scale variables, 1 = never use L1/L2 with pets, 5 = daily use L1/L2 with pets)

The PCA employed Varimax rotation, listwise deletion of missing values, and components were selected based on Eigenvalues > 1. The rotated component matrix, showing full component loading after Varimax rotation,[5] is listed in the Online Appendix, Table A2 (loadings greater than .4 are considered significant). Six components were extracted based on this matrix, accounting for a total of 73.03% of variance.

The outcomes of the PCA fall into distinct categories, relating to language proficiency (Component 1 and 4) on the one hand, and language experience (Component 2, 3, 5, and 6) on the other. In several cases, measures relating to both L1 and L2 loaded onto the same component, suggesting that responses or scores for the two languages are not independent

The first component appeared to relate mainly to *L2 proficiency*, comprising high loadings for the two L2 proficiency measures (C-Test and Can-Do scale) as well as for the holistic self-assessment question. Interestingly, the L1 C-Test also loaded highly on this component. The second component related to the use of the L1 outside the home, i.e., with friends and at work, while the fourth component comprised the *language use with relatives*, with the L1 loading positively and the L2 negatively on this component,[6] The third component represented *perceived L1 proficiency*, comprising the holistic self-assessment question, the Can-Do scale, and the language preference question. The fifth component relates to the *use of the L2 with friends* and is composed of the frequency of L2 use with friends and the question of what the native language of most of the social contacts is. The final component was *language use with pets* (this did not emerge as a significant predictor in any of the preliminary models and was therefore not included in any of the analyses presented below).

A preliminary group comparison by means of one-way ANOVA with Tukey Post-Hoc tests showed a main effect for Group on all five components (see Appendix Table A3 for details). With respect to L2 proficiency, the German native speakers outperformed the groups with Spanish as either their L1 or their L2, indicating that their proficiency in L2 English was higher than that of the Spanish natives and higher than the Spanish proficiency of the English natives. There was no difference between the L2 proficiency of either German L1 vs. German L2 or Spanish L1 vs. Spanish L2; nor between English natives German vs. Spanish L2 (see Figure 1).

[5]Varimax rotation is an orthogonal rotation method used to maximize the variance of squared loadings within each factor. This results in a simpler and more interpretable factor structure by enhancing the distinction between factors, making it easier to identify which variables strongly associate with each component. Unlike oblique rotations, Varimax assumes that factors remain uncorrelated, which is helpful when analyzing results through multiple regression, as we do here.

[6]The fact that this component includes both the L1 and the L2, with opposite polarities, may seem surprising, but is a reflection of the fact that language use with relatives exists in a circumscribed setting—unlike with friends, having more speakers of one language within the family means having fewer of the other, e.g., when a participant's partner, and consequently their in-laws, are L2 speakers.

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

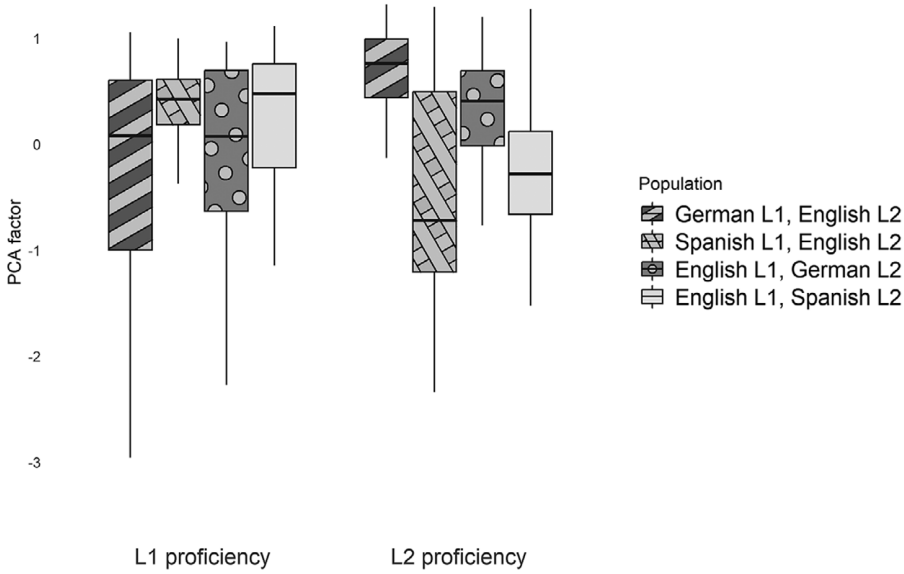


Figure 1. Proficiency in the first and the second language across bilingual groups.

The pattern of frequency of use varied across domains (see Fig. 2): the German L1 speakers seemed to use the L1 less within the family domain than the other groups (who did not differ from each other). The English L1–German L2 speakers had a higher tendency to use their German L2 with friends and at work, while for the native Spanish speakers with English as their L2, there was comparatively little use of the L2 with friends.

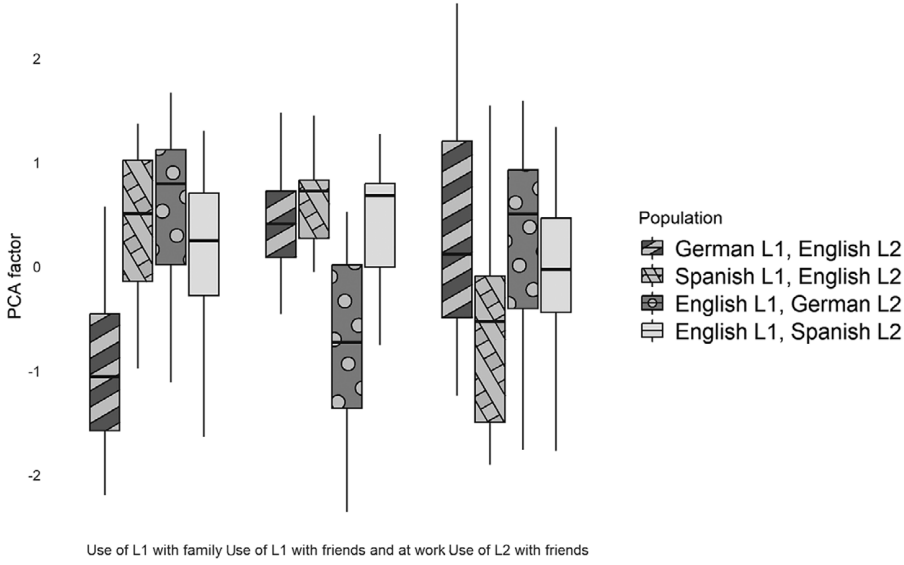


Figure 2. Use of the L1 and the L2 across domains by bilingual populations.

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press



Figure 3. Example stimulus from Picture Naming Task.

The descriptive and inferential statistics for the predictors extracted by the PCA are given in the Online Appendix, Table A4.

Lexical access tasks

Picture naming task

The naming experiment was conducted in person in a quiet environment, with only the participant and the experimenter present.[7] All participants completed a timed Picture Naming Task (PNT) in their L1; bilingual participants also completed the task in their L2. Presentation of items was pseudo-randomized into two lists (counterbalanced), to ensure that no target item was semantically or phonologically similar to the preceding trial.

We selected 40 line drawings for the Spanish version of the task and 60 items for the German version[8] from the database provided by the International Picture Naming Project (IPNP, Bates et al., 2003), see Figure 3 for an example (the list of items is given in the Online Appendix, Table A5).

The IPNP is a freely available database, based on a comparison of 512 line drawings of objects across seven languages in terms of a broad range of factors, including frequency, naming speed, and recognizability Matching lexical frequencies across languages is complex, as frequency counts depend on the size and characteristics of the corpora they were derived from,[9] and Bates et al. propose a "universal frequency" factor, derived through a principal component analysis of a range of frequency measures for the seven languages (see Bates et al., 2003, for details).

[7]The experiment also comprised a Flanker test of inhibitory control. Due to equipment failure during the fieldwork conducted in Spain, not all participants were able to complete this task. We are not reporting it here.

[8]The experiment was part of a wider study not reported here intended, which investigates grammatical gender. The Spanish list thus included 20 masculine and 20 feminine nouns, while the German list included 20 items each for masculine, feminine, and neuter nouns.

[9]The IPNP Bates et al. (2003) used the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995) for English and German frequencies and Alameda & Cuetos (1995) for Spanish frequency.

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

The items chosen for the present experiment were spread across this universal frequency measure (mean .31, SD 1.07, range −1.36 to 2.93). This measure correlated with the language-specific log-transformed frequency measure for these items reported by Bates et al. (2003) at above .8 (English: .89, Spanish: .83, German: .87), and there was no significant difference between the frequency factor of the items included in each language ($F(1, 177) = .42$, $p = .518$; for full descriptive statistics see Appendix, Table A6). Word length (in characters) did not differ between languages (English: 5.29, German: 5.73, Spanish: 5.75, $F(1, 177) = 2.033$, $p = .16$; for full descriptive statistics see Appendix, Table A6). We had initially hoped to be able to exclude cognate items, but this proved to be impossible, and cognateness is therefore included as a dichotomous factor in the statistical models. Items were coded as cognates if the Spanish and German words shared clear phonological and orthographic similarity stemming from common etymology (e.g., Spanish *león*/*lion* or German *Anker*/*anchor*). Items were coded as non-cognates if the words were etymologically unrelated or showed substantial phonological divergence (e.g., *manzana/apple; Hose/trousers*). The resulting lists of items to be named had an average frequency value of .44 for the German version of the task and .30 for the Spanish version (the difference was not significant, $t(101) = .633$; $p = .528$), with 28 items in the German version and 15 in the Spanish one being cognates with English.

The pictures were presented on a 17-inch laptop screen with E-Prime version 2.0. Each trial started with a 100 ms beep and a fixation cross displayed in the middle of the screen for 500 ms, with the picture displayed immediately afterward. The display lasted for up to 4000 ms or until a verbal response was detected by the VoiceKey microphone. Each experiment began with eight practice items, after which the participant was given the opportunity to ask questions before proceeding to the main experimental block.

Picture-word interference task

In addition to the PNT, bilingual participants completed a Picture-Word Interference Task (PWIT) in both their languages and monolinguals in their L1. The task used the same stimulus items as the PNT, but a distractor word was added across the picture in sans-serif capitals in the language of the task (see Figure 4). The order of tasks was counterbalanced across participants.

Distractor items were semantically and phonetically unrelated to the target item. In all other respects, the experiment was identical to the procedure described under section "Picture naming task." Presentation of the type (PNT vs. PWIT) and language (L1 vs. L2) of the task was counterbalanced across participants.

Repetition of items

All experiments used the same list of items, so that bilingual participants encountered each picture four times while monolinguals encountered them twice. For half of the participants, the first (or only) PNT (List 1) preceded the first (or only) PWIT (List 2); for the other half, it was the other way around. This order was reversed for each bilingual participant in the second session in the other language. Bilinguals attended two sessions: in the first, they were tested in German and Spanish; in the second (between 2 weeks and 4 months later), they were tested in English.

We opted for this design, as it allowed us to directly compare the impact of both the target language and the presence of the distractor item for each participant on an item-base (see below, Section "Results"); and this choice inevitably led to the repetition of

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press



Figure 4. Example stimulus from Picture-Word Interference Task.

pictures for each participant. Previous research has consistently shown repetition priming across short periods of time (several minutes) (e.g., Francis, Augustini, & Sáenz, 2003; Gollan et al., 2005; Ivanova & Costa, 2008). Repetition effects have been shown to be robust across item frequency (Gollan et al., 2005; Ivanova & Costa, 2008), but differences between monolinguals and bilinguals are not consistently affected across studies: while Gollan et al. (2005) found that the difference between monolinguals and bilinguals disappeared from the fourth repetition, Ivanova and Costa (2008) were unable to replicate this leveling.

We attempted to minimize the impact of repetition priming as far as possible within our design: each session administered the PNT and the PWIT in the same language, but counterbalanced the order—half of the participants encountered the PNT first, the other started with the PWIT. Several other tasks not reported here were completed between the two tasks, so that the second task typically started 20–30 minutes after the first. The English experiments were conducted in a separate session—due to participant and lab availability, the interval could not be kept constant, but the minimum time between sessions was 2 weeks.

In order to compensate for the effect of repetition priming, the order of tasks is furthermore included in our statistical models.

Data processing

All experiments were audio recorded on a Tascam recorder. Response latencies were subsequently manually measured in Praat (Boersma & Weenink, 2024) by marking the onset of both the beep preceding the trial and the verbal response and then subtracting 500 ms (the time between the onset of beep/fixation cross and image) from the resulting interval. Responses were coded for accuracy, and inaccurate or empty responses, as well as responses preceded by a hesitation or false start (4.59%), were excluded from the analyses. It is customary in naming tasks to exclude very fast responses, which are unlikely to reflect complete cognitive processing and may represent anticipatory or accidental responses, as well as slow responses, which likely reflect lapses in attention or task disengagement rather than genuine naming latency. While there is no commonly agreed threshold for either of these cutoffs, the lower threshold is usually between 150

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

and 300 ms (e.g., Nappo et al., 2022; Botezatu et al., 2022). Since the data had been hand-coded to eliminate false triggers, we opted for a lower threshold at the higher end of the range and excluded 231 responses below 250 ms (1.48%). For the upper threshold, we identified outliers by first trimming the data to exclude 234 responses >4000 ms in order to avoid extreme times distorting the calculation. RTs above the mean plus two standard deviations of the trimmed data per language were then also excluded (n = 519, 3.33%).

Split-half reliability was assessed by randomly dividing the target items into two halves across 100 iterations, calculating mean reaction times for each participant in each half, and correlating these values with the Spearman-Brown correction applied. All six Task × Language combinations demonstrated excellent internal consistency (mean corrected r > .90, see Figure 1, Appendix), indicating that participants' reaction time patterns were highly stable across different item sets.

Results

In order to answer our Research Questions 1–3, we derived three different outcome variables from the PNT and the PWIT, respectively. The first variable was the log10-transformed naming latency, to determine absolute naming speed (RQ1). Secondly, we attempted to factor out response speed as an individual characteristic in order to measure dominance (RQ2). This was achieved by subtracting the log10-transformed response latency that the bilingual participants had for the same item on the same task in their L1 from that of their L2. So, if the item was, for example, "apple," a positive value would indicate that a Spanish native speaker had been faster in saying *manzana* in the Spanish version of this task than *apple* in the English version (while a negative value would indicate a faster response in the L2 than in the L1), and vice versa for an English-Spanish bilingual. Finally, in order to assess how the necessity to inhibit the unrelated distractor item would impact response latencies (RQ3), we subtracted the log10-transformed item response latency in the PNT from that of the PWIT separately for each item in each of the participants' languages, with a positive value indicating the participant had responded faster to the same item in the same language when there was no distractor than when there was. While all statistical models are based on the log10-transformed RTs, we used the millisecond measures for any visualizations, for ease of interpretation.

The descriptive statistics for these three variables are summarized in Tables 4 (RT and PWIT cost) and 5 (difference L1-L2). For ease of interpretation, these tables give the base values in place of the log-transformed RTs which we use in the models.

We built three sets of Linear Mixed-Effects Regression (LMER) models, one for each of these different outcome variables. Since the experiment comprised three target languages (English, German and Spanish), with all bilingual participants completing half of the naming tasks in English, but being split into a Spanish and a German subgroup for the other two, we created nested contrasts to be able to represent this in the models. The first contrast compared the target language English with the two other target languages (Spanish and German) combined, whereas the second contrast compared German with Spanish. In addition, we created sum-coded contrasts (−.5, .5) to represent the binary variables of cognate status (cognate vs. non-cognate), Task (PNT vs. PWIT), accuracy of the previous item (accurate vs. inaccurate—inaccurate items were not included in the analysis; but there may be a delay in naming latency for an accurately named item following an error or failure to retrieve the previous item, and this delay may

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

Table 4. Average Naming Speed and PWIT Cost (ms) by Population

		RTs (ms)				PWIT cost	
		PNT		PWIT		PWIT–PNT	
	Language of task	mean	std	mean	std	mean	std
GEMono	German	706	255	781	261	79.73	299.57
GEL1ENL2	L1 (German)	789	280	895	279	108.13	343.38
	L2 (English)	791	313	877	282	69.80	327.40
ENL1GEL2	L1 (English)	735	280	812	281	81.40	299.18
	L2 (German)	926	334	945	305	7.07	405.14
SPMono	Spanish	709	289	747	314	43.51	373.55
SPL1ENL2	L1 (Spanish)	767	281	823	269	62.98	343.66
	L2 (English)	863	306	883	317	85.80	452.61
ENL1SPL2	L1 (English)	709	253	798	279	110.79	326.99
	L2 (Spanish)	830	315	926	307	107.23	366.06
ENMono	English	819	430	770	296	–83.87	504.65

Table 5. Average Difference Measure in Naming Speed (ms) Between L1 and L2 by Population

	PNT		PWIT	
	mean	std	mean	std
GEL1_ENL2	49.86	339.60	–2.11	334.66
ENL1_GEL2	232.33	369.40	161.00	355.19
SPL1_ENL2	86.67	323.61	121.35	443.88
ENL1_SPL2	189.64	379.15	178.30	407.28

distort overall reaction times for bilinguals in their weaker language), and whether the language of the task was the participant's L1 or L2. "List," indicating the order of experiments (PNT first or PWIT first), was similarly sum-coded (−.5, .5). All models were built in R version 4.3.0 (R Core Team, 2023) with the lme4 package (Bates, Maechler, Bolker, & Walker, 2015).

Log10-transformed RT in both tasks and both languages (Analysis I)

Our first analysis investigated RQ1 and the hypothesis that monolinguals would be faster, in absolute terms, than bilinguals when naming items, and that bilinguals would name them faster in their L1 than in their L2. Our initial model thus included the log10-transformed RT from all participants, all tasks, and all languages. (All full model specifications are available in the Appendix, an overview summary is provided in Table 6.)

Predictors were entered in two steps: the first step comprised predictors related to the experimental design: *frequency of items, word length (number of characters), cognateness*, *accuracy of the previous item*, *task* (PNT vs. PWIT), and *item repetition* (List 1 vs. List 2). It furthermore included the contrast-coded variables *language of the task* and *Group*, as well as the *age* and *education level* of the participants.

Word length, age, and education were not significant predictors and were therefore not included in the further analyses. All other predictors were highly significant: participants were faster to name more frequent (estimate = −0.025, SE = 0.001, t =

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

−28.682, p < .001) and cognate (estimate = 0.036, SE = 0.002, t = 19.249, p < .001) items, and errors on an item caused a delay in the naming of the next one (estimate = −0.039, SE = 0.005, t = −8.454, p < .001) (recall that inaccurate items themselves were not included in the analyses). Items were named faster in the PNT than in the PWIT (estimate = −0.037, SE = 0.002, t = −22.069, p < .001), and faster when they were encountered for the second time within the same session (estimate = −0.020, SE = 0.001, t = −16.585, p < .001). Task language was significant, with responses in English faster than in the other languages (combined) (estimate = −0.018, SE = 0.002, t = −10.371, p < .001) and Spanish responses faster than German responses (estimate = 0.037, SE = 0.005, t = 7.476, p < .001). Bilinguals performed more slowly than monolinguals (estimate = −0.048, SE = 0.006, t = −7.376, p < .001) with speakers with English L1 outperformed by speakers with German or Spanish L1 (estimate = −0.042, SE = 0.008, t = −5.503, p < .001). An interaction term between frequency and group showed that the frequency effect was somewhat less pronounced in the monolingual groups than in the bilinguals (estimate = 0.002, SE = 0.001, t = 2.205, p < .05), while bilinguals with English as their L1 did not differ from bilinguals with German or Spanish as their L1 (estimate = 0.000, SE = 0.002, t = −0.034, p = 0.973). The total explained variance was .33.

In the second step, we reduced the dataset to comprise only bilingual participants, in order to be able to assess the impact of the bilingualism-specific predictors, retaining all predictors that had been significant in the previous step. We first determined to what extent participants' response times would vary depending on whether they were naming an item in their L1 or their L2, and found that participants were faster in their L1 than in their L2 (estimate = 0.02, SE = 0.00, t = 7.300, p < .001). An interaction term between language of the task and item frequency was significant, indicating that frequency had a stronger effect in the L2 than in the L1 (estimate = −0.00, SE = 0.00, t = −2.554, p < .05). We also found a significant interaction between the group (L1 English vs. L2 English) and whether the task was in the L1 or in the L2: the immersed bilinguals for whom English was the second language were virtually equally fast in the L1 and the L2, whereas the non-immersed bilinguals were slower in their L2 (estimate = 0.06, SE = 0.00, t = 11.822, p < .001).

Finally, we added the five components that were derived from the Language Background Survey by means of the PCA (see above, section "Participants") to the model, in each case calculating the interaction term with whether the task was completed in the participant's L1 or L2. There was no interaction between L1 proficiency and the language of the task, but a highly significant interaction term between L2 proficiency and the language of the task (estimate = −0.01, SE = 0.00, t = −4.476, p < .001), indicating that participants with higher L2 proficiency were faster to name items in their L2, but no such effect was found for the L1. None of the measures of frequency of L1 use were significant, either on their own or in interaction with the language of the task. The total explained variance was .36.

Difference in naming latency per item in L1 vs. L2 (Analysis II)

The second set of analyses investigates language dominance among the bilingual populations, attempting to establish how the factors investigated in relation to absolute naming times above may modulate response latencies in the L1 vs. the L2—that is, how language dominance is affected. Since Analysis II was based on a variable that represented the difference in naming latency for each item in the participants' L2 vs. their L1 (e.g., speed of saying *apple* minus speed of saying *manzana*), the model was

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

Table 6. Summary of Results

Factor		finding	Analysis I—log RT	Analysis II—RT L2–RT L1	Analysis III—RT PNT–RT PWIT
	frequency	high frequency < low frequency	***	*	*
item level	word length	longer words < slower RT	n.s.		
	cognateness	cognate < non-cognate	***	**	n.s.
	accuracy	previous item accurate < inaccurate	***		n.s.
task level	task	PNT < PWIT	***	***	
	order	first encounter > next encounter	***		
	language	English < German/Spanish	***	***	n.s.
		Spanish < German	***	n.s.	n.s.
population level	Group	monolinguals < bilinguals	***		n.s.
		English L1 > German/Spanish L1	***	***	n.s.
	L1 vs. L2	L1 < L2	***		***
	frequency * Group	monolinguals less affected by item frequency than bilinguals	***		n.s.
interactions	frequency * L1 vs. L2	frequency effect more pronounced in L2 than in L1 of bilinguals	*		
	Group * L1 vs. L2	gap between L1 and L2 RTs smaller for immersed than non-immersed	***		
	age		n.s.	n.s.	n.s.
	education		n.s.	n.s.	n.s.
	L2 proficiency * L1 vs. L2	higher L2 proficiency: faster L2 naming	***	(*)	n.s.
individual level	L1 proficiency		n.s.	n.s.	n.s.
	Language use with family		n.s.	n.s.	n.s.
	L1 use with friends/at work		n.s.	n.s.	n.s.
	L2 use with friends		n.s.	n.s.	n.s.
total R^2			.36	.21	.21

*: p < .05, **: p < .01, ***: p < .001, (*) = p <.1.

built using a subset of the data comprising the bilingual participants only, and the language in which the task was conducted or whether it was the participant's L1 or L2 were not applicable predictors.

The gap between the L1 and the L2 was smaller for more frequent (estimate = −.007, SE = .003, t = 2.437, p < .05) and for cognate (estimate = .017, SE = .006, t = 2.812, p < .01) items. It was larger for the PNT than the PWIT (estimate = .031, SE = .005, t = 6.772, p < .001) and larger for English natives than for English L2 speakers (estimate = .040, SE = .012, t = 3.232, p < .001), but Spanish and German speakers did not differ from each other (estimate = −.013, SE = .023, t = −.581, p = .563). Age and education were again not significant predictors. None of the PCA components significantly impacted this measure of language dominance; the only marginally significant effect was that a higher level of L2 proficiency was associated with a smaller gap between L1 and L2 (estimate = −0.018, SE = 0.010, t = −1.879, p = .065). None of the measures relating to frequency of use or L1 proficiency emerged as significant. The total explained variance was .21.

Impact of picture-word interference paradigm (Analysis III)

In the final set of analyses, we investigated the impact of the presence of a (semantically and phonologically unrelated) distractor item printed across the picture. In order to calculate the effect of this manipulation, we subtracted the naming latency with which each participant had named an individual item in the PNT from that for the same item in the PWIT in the same language. The delay caused by the distractor was larger for more frequent items (estimate = .007, .003, t = 2.273, p < .05) but unaffected by cognate status, accuracy of the previously named item, item repetition, age, or education. There was no difference in the impact the presence of a distractor item had for monolinguals vs. bilinguals, or for the different bilingual subpopulations, but the gap was larger in participants' L1 than in their L2 (estimate = .017, .004, t = 4.103, p < .001).

We again continued the analyses with a subset of the data comprising the bilingual participants. The analyses conducted on this subset only found a significant gap between participants' L1 and their L2 (estimate = .019, .004, t = 4.649, p < .001). None of the composite scores derived by the PCA had a significant effect on the delay caused by the presence of a distractor item. The total effect size was .21.

Discussion

In this study, we assessed the speed of lexical naming in bilingual and monolingual populations. While the monolinguals only named items in their L1, bilinguals conducted the experiment in two sessions, one in their L1 and one in their L2. The items to be named were spread across the frequency range, from high-frequency items such as *hand* or *house* to low-frequency ones (*rake*, *asparagus*) (Bates et al., 2003). We assessed a range of variables capturing the background, language habits, and proficiency of our participants: two tasks (a C-Test and a Can-Do scale) and various self-assessment questions were used to estimate L1 and L2 proficiency, and frequency of use of either language in a range of contexts. From these variables, we derived five composite descriptors of L1 and L2 proficiency and use.

We initially investigated the impact of factors relating to the stimuli or trials: item frequency, word length, cognateness, and recovery time after an item was named incorrectly or not retrieved. Word length was not a significant predictor of response

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

speed, but our analyses confirmed previous findings for the other three factors: All groups named more frequent items faster than less frequent ones, and the slope associated with frequency was steeper in the second than in the first language and steeper for bilinguals than for monolinguals. Cognate items were named faster than noncognates, and the cognate facilitation effect was stronger in the L2 than in the L1. Errors or retrieval failure had a "spillover" effect on the next item to be named, and this recovery time was longer in participants' L2 than in their L1. Of these factors, word frequency was the only one to impact the difference between response times in the presence vs. absence of a distractor item.

RQ1—Lexical access: In the next step, we assessed characteristics of the populations. Our first research question related to the differences in naming speed between the populations. We had two hypotheses here, namely, that bilinguals would a) perform slower than monolinguals in both their L1 and their L2, and b) perform slower in their L2 than in their L1, and both of these hypotheses were confirmed by Analysis I. There are three possible, and not mutually exclusive, explanations for this finding:

a) naming in bilinguals is delayed due to a larger repertoire (e.g., Ramscar et al., 2013)
b) naming in bilinguals is delayed due to a lower frequency with which items are encountered (e.g., Gollan et al., 2005; 2008)
c) naming in bilinguals is delayed due to the additional task of inhibiting competitors in the other language (e.g., Shao, Meyer, & Roelofs, 2013)

In order to assess to what extent these three explanations contribute to overall speed of lexical access for our participants, we included a range of predictors. Language proficiency was included on the assumption that a higher level of proficiency is a proxy of a larger repertoire (note that we did not include any tasks explicitly measuring vocabulary size). The potential impact of this variable is complex: higher L2 proficiency can be assumed to facilitate naming in the L2; however, if a larger repertoire leads to overall slower naming, higher L2 proficiency, which implicates a larger vocabulary, should slow down naming in the L1. This was not the case, with L2 proficiency indeed associated with faster L2 naming, but not with any delays in the L1.

The second potential explanation, namely that bilinguals are delayed due to the fact that they have to spread their time across their languages, is more complex. Our bilingual participants performed slower overall than the monolinguals, with the immersed participants (who presumably spend more time on their L2 and less on their L1) not only outperforming the non-immersed ones, but being almost equally fast in their L1 and their L2, while the non-immersed bilinguals performed much slower in L2 than in L1—however, both groups were almost equally fast in their L1. This finding implies that a higher frequency of use of a second language learned later in life allows RTs to speed up until they eventually catch up with the naming latencies of L1 items. On the other hand, a decrease in use of the L1 does not appear to cause a similarly substantive slowdown of naming speed. It is also interesting that, while the overall group effect demonstrates the differential effect of immersed vs. non-immersed bilingualism, the individual measures on reported usage of L1 and L2 across a range of contexts do not appear to play a role.

We will return to the third explanation listed here below, when we discuss RQ3.

RQ2—Language dominance: The second research question related to the relative speed with which individual bilinguals would name items in their L1 vs. their L2. This analysis was done in order to factor out the individual variable of overall response speed and focus on how much faster or slower each participant had named the same item in

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

their L2 vs. their L1, on the assumption that naming in the dominant language would be faster. Two hypotheses were formulated here: firstly, that the difference measure would be associated with L1 and L2 proficiency, and secondly, that it would be modulated by individual frequency of exposure, with more use of the L1 associated with a larger gap, and more use of the L2 with a smaller one.

The descriptive statistics show that, out of the four bilingual groups, only the German speakers in the UK performed similarly fast in their L1 and their L2 at group level, and they were also the only population for whom the proficiency tests indicated a relatively balanced level; all other populations were faster at naming in their L1 and scored higher on the L1 proficiency tasks than in their L2. The statistical analyses confirmed Hypothesis IIa, in that higher L2 proficiency was associated with a marginally smaller difference between L2 and L1. With respect to frequency of use, we again found a group effect, showing a smaller gap between L1 and L2 response speed for the immersed bilinguals than for the non-immersed population, but no impact of the individual measures of differential frequency of use across a variety of domains.

RQ3—Inhibition: As predicted, all populations were slower to name items in the PWIT than in the PNT. We had hypothesized that the additional effort of selectively inhibiting an unrelated competitor item in the PWIT, which has been shown to incur demands similar to those of inhibiting competitors from the non-target language, would lead to a larger gap between PNT than PWIT for bilinguals (where the two inhibitory efforts compound each other). This hypothesis was confirmed. Unlike in the other analyses, there was no difference between the immersed and the non-immersed populations, but the gap between PWIT and PNT was larger in participants' L2 than in their L1.

RQ4—Individual differences: Our final research question and set of hypotheses addressed the question of how item characteristics and individual characteristics would conspire to affect individual naming speed. Here, however, we drew a blank: while our analyses confirmed previous findings that item frequency would have a stronger impact on naming speed in the L2 than in the L1, and while there was some support for the assumption that immersed populations would have stronger L2 performance than non-immersed participants, we could not find any evidence of interactions between this factor and either proficiency or frequency of use in either language.

There are some interesting insights to be drawn from these findings beyond the fact that they confirm previous investigations with respect to the impact of frequency,

Table 7. Summary of RQs, Hypotheses, and Findings

RQ1—lexical access	
HIa: monolinguals < bilinguals	confirmed
HIb: bilinguals in L1 < bilinguals in L2	confirmed
RQ2—language dominance	
HIIa: higher L2 proficiency ≥ smaller gap between L1 and L2	confirmed
HIIb: frequency of exposure associated with dominance measure	confirmed for group (immersed vs. non-immersed), not confirmed on individual level
RQ3—inhibition	
HIII: larger PWIT cost for bilinguals	confirmed
RQ4—individual differences	
HIVa: interaction item frequency and frequency of use	not confirmed
HIVb: interaction item frequency and proficiency	not confirmed

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

cognateness, and bilingualism vs. monolingualism. While our predictions relating to the level of the population are largely borne out by our analyses, at the individual level, we find much less support for our assumptions. The exception here is the relatively straightforward measure of L2 proficiency, which has been shown to facilitate naming in the L2, leading to more balanced performance in both languages.

Much more confusing is the factor of language use. This factor poses a perennial puzzle in studies of L1 attrition, as has often been pointed out in the past (see Schmid, 2019, for an overview), and it does not disappoint here, either. Theories of entrenchment or activation thresholds predict a straightforward association, with more frequent use facilitating both access and accuracy, but this association has never been confirmed for the L1 of bilinguals, and it is not confirmed here, either. While we find that immersed speakers overall become faster in their L2, neither immersion vs. non-immersion nor any of the measures of individual use seem to facilitate or hinder L1 response latency. One possible explanation is that self-report data, on which the assessment of frequency of use in long-term bilinguals has to rely, may not be fine-grained or accurate enough—however, it is striking that a broad range of studies using different approaches to the measurement of bilingual language exposure and use have, to date, not been able to find a correspondence between these measures and any aspect of L1 performance in an attrition setting. Schmid (2007) proposes that late bilinguals, who use one language exclusively or virtually exclusively until they reach puberty, may reach a "saturation threshold" of activation levels which remain impervious to a decline in use and activation later on, and this seems to be what we have found here, as well.

Conclusion

This study set out to explore the impact of frequency as a predictor of naming speed in bilinguals, and in particular to explore interactions between the item-level variable of word frequency and the individual characteristic of frequency of language use in bilinguals' first and second languages. While our analyses clearly confirmed earlier findings that higher-frequency items are named faster in both languages, with a steeper slope in participants' L2 than in their L1, and while it seems that immersed bilinguals have an advantage over non-immersed ones with respect to naming speed in the L2, we did not find any suggestion that naming speed in the L1 is similarly affected by the linguistic environment, and neither L1 nor L2 naming was affected by self-reported individual levels of L1 or L2 use across a range of domains. This finding throws into question some widely held notions relating to the activation and entrenchment of linguistic items. We found that generally, even long-term immersed late bilinguals with high levels of L2 proficiency and frequent L2 use do not become faster at naming in the L2 than in the L1 (at most, some balanced bilinguals attain similar reaction times, but overall, the L1 remains faster). Particularly intriguing is the finding that bilinguals do appear to become slower in their L1 than monolinguals of the same language, but that this effect appears to be the outcome of adding an additional language, rather than of a decrease of exposure in that language. The slowdown in lexical retrieval observed in the L1 of bilinguals thus appears to be the outcome of the added cognitive pressure of managing more than one linguistic system, rather than of a weakening of entrenchment due to a decline in exposure.

The present study had some limitations, most obviously that our intended participant pool could not be attained, due to the pandemic. It is possible that the failure of some of the individual predictors to have a significant effect is due to a lack of power. Future work should aim to increase population sizes to assess this.

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0272263126101806.

Data availability statement. The anonymised data and scripts for the analyses reported here are available on OSF: https://osf.io/38b2j/overview?view_only=8f4c607d136a468e996a6950d29f5fab.

Acknowledgments. The study presented here was supported by a Research Priming fund from the University of Essex.

Competing interests. The authors declare none.

References

Alameda, J. R., & Cuetos Vega, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Universidad de Oviedo, Servicio de Publicaciones.

Alario, F. X., Costa, A., & Caramazza, A. (2002). Frequency effects in noun phrase production: Implications for models of lexical access. *Language and Cognitive Processes*, *17*(3), 299–319.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (release 2). Distributed by the linguistic data consortium*, University of Pennsylvania.

Bates, E., D'amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C. C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., … & Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*, 344–380.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Baus, C., Costa, A., & Carreiras, M. (2013). On the effects of second language immersion on first language production. *Acta Psychologica*, *142*(3), 402–409.

Bergmann, C., Sprenger, S. A., & Schmid, M. S. (2015). The impact of language co-activation on L1 and L2 speech fluency. *Acta psychologica*, *161*, 25–35.

Boersma, P. & Weenink, D. (2024). Praat: doing phonetics by computer [Computer program]. Version 6.4.07, retrieved 17 March 2024 from http://www.praat.org/

Botezatu, M. R., Misra, M., & Kroll, J. F. (2022). Proficiency in a second language influences processing of print-to-sound mappings. *Linguistic Approaches to Bilingualism*, *14*(3), 285–309.

de Bruin, A., Della Sala, S., & Bak, T. (2016). The effects of language use on lexical processing in bilinguals. *Language, Cognition and Neuroscience*, *31*(8), 967–974. https://doi.org/10.1080/23273798.2016.1190024.

de Jong, N. H., Groenhout, R., Schoonen, R., & Hulstijn, J. H. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, *36*(2), 223–243.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and language*, *62*(3), 284–301.

Gollan, T. H., & Acenas, L. A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(1), 246.

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. (2005). Bilingualism affects picture naming but not picture classification. *Memory & cognition*, *33*, 1220–1234.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, *1*(2), 67–81.

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530.

Hulsen, M. (2000). *Language loss and language processing: Three generations of Dutch migrants in New Zealand*. PhD thesis, Katholieke Universiteit Nijmegen.

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production?. *Acta psychologica*, *127*(2), 277–288.

Jarvis, S. (2019). Lexical attrition. In M. S. Schmid and B. Köpke (Eds.), *The Oxford Handbook of Language Attrition* (241–250). Oxford University Press.

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.

Kastenbaum, J., Bedore, L., Peña, E., Sheng, L., Mavis, I., Sebastian-Vaytadden, R., … Kiran, S. (2019). The influence of proficiency and language combination on bilingual lexical access. *Bilingualism: Language and Cognition*, *22*(2), 300–330. https://doi.org/10.1017/S1366728918000366

Köpke, B., & Genevska-Hanke, D. (2018). First language attrition and dominance: Same same or different?. *Frontiers in Psychology*, *9*, 397002.

Langacker, R.W. (1987). *Foundations of Cognitive Grammar. Vol I: Theoretical prerequisites*. Stanford University Press.

Lehtinen, N., Kautto, A., & Renvall, K. (2024). Frequent native language use supports phonemic and semantic verbal fluency in L1 and L2: An extended analysis of verbal fluency task performance in an L1 language attrition population. *International Journal of Bilingualism*, *28*(5), 884–906.

Linck, J. A., Kroll, J. F., & Sunderman, G. (2009). Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological Science*, *20*(12), 1507–1515.

Nappo, R., Galati, G., Bureca, I., & Romani, C. (2022). Semantic interference and facilitation in picture naming: The effects of type of impairment and compensatory strategies. *Cognitive Neuropsychology*, *39*(5-8), 325–355.

Neumann, E. (2018). Bilingual cross-language priming reveals common inhibitory modulation effects in selective attention and memory. *Journal of Brain Behaviour and Cognitive Sciences 1*(2:9), e3.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*(4), 273–281.

Opitz, C. (2010). L1 attrition and L2 acquisition: Global language proficiency and language dominance in adult bilinguals. *Eurosla Yearbook*, *10*(1), 248–281.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.

Paradis, M. (2007). L1 attrition features predicted by a neurolinguistic theory of bilingualism. In B. Köpke, M. S. Schmid, M. Keijzer, & S. Dostert (Eds.), *Language Attrition: Theoretical perspectives* (pp. 121–134). John Benjamins.

Ramscar, M., Hendrix, P., Love, B., & Baayen, R. H. (2013). Learning is not decline: The mental lexicon as a window into cognition across the lifespan. *The Mental Lexicon*, *8*(3), 450–481.

Schmid, H.-J. (2017). A framework for understanding linguistic entrenchment and its psychological foundations. In H.-J. Schmid (Ed.), *Entrenchment and the psychology of language learning* (pp. 9–36). Mouton de Gruyter.

Schmid, M. S. (2007). The role of L1 use for L1 attrition. In B. Köpke, M. S. Schmid, M. Keijzer & S. Dostert (Eds.), *Language attrition: Theoretical perspectives* (pp. 135–153). John Benjamins.

Schmid, M. S. (2019). The impact of frequency of use and length of residence on L1 attrition. In M. S. Schmid and B. Köpke (Eds.), *The Oxford Handbook of Language Attrition* (pp. 288–303). Oxford University Press.

Schmid, M. S., & Beers Fägersten, K. (2010). Disfluency markers in L1 attrition. *Language Learning*, *60*(4), 753–791.

Schmid, M. S., & Yılmaz, G. (2021). Lexical access in L1 attrition—competition versus frequency: A comparison of Turkish and Moroccan attriters in the Netherlands. *Applied Linguistics*, *42*(5), 878–904.

Shao, Z., Meyer, A. S., & Roelofs, A. (2013). Selective and nonselective inhibition of competitors in picture naming. *Memory & Cognition*, *41*, 1200–1211.

Shao, Z., Roelofs, A., Martin, R. C., & Meyer, A. S. (2015). Selective inhibition and naming performance in semantic blocking, picture-word interference, and color–word Stroop tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1806.

Shook, A., Goldrick, M., Engstler, C., & Marian, V. (2015). Bilinguals show weaker lexical access during spoken sentence comprehension. *Journal of Psycholinguistic Research*, *44*(6), 789–802.

Starreveld, P. A., & La Heij, W. (2017). Picture-word interference is a Stroop effect: A theoretical analysis and new empirical findings. *Psychonomic Bulletin & Review*, *24*, 721–733.

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, *20*(4), 912–928.

Sudarshan, A., & Baum, S. R. (2019). Bilingual lexical access: A dynamic operation modulated by word-status and individual differences in inhibitory control. *Bilingualism:* Language and Cognition, *22*(3), 537–554.

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.

Yilmaz, G., & Schmid, M. S. (2012). L1 accessibility among Turkish-Dutch bilinguals. *The Mental Lexicon*, *7*(3), 249–274.

Cite this article: Schmid, M., Martin Gascon, B., Soto, C., & Solaimani, E. (2026). Anything you can say, I can say faster: The impact of item frequency, frequency of language use, and inhibitory control on monolingual and bilingual picture naming speed. *Studies in Second Language Acquisition,* 48: 784–807. https://doi.org/10.1017/S0272263126101806

https://doi.org/10.1017/S0272263126101806 Published online by Cambridge University Press