MASTERARBEIT | MASTER'S THESIS

Titel | Title

Exploring Working Memory in a Highly Multilingual Sample: The Degree of Multilingualism, Formality of Language Acquisition and Individual Variability in Working Memory

verfasst von | submitted by

Amin Mohammed BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 840
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Psychologie
Betreut von | Supervisor:	Assoz. Prof. Narly Golestani BSc PhD
Mitbetreut von | Co-Supervisor:	lic. Olga Kepinska MA MA PhD

Abstract

Hintergrund und Zielsetzung. Positive Effekte der Zweisprachigkeit auf exekutive Funktionen (EF) sind umstritten (z. B. Antoniou, 2019), da Metaanalysen keine bis geringe Effekte aufzeigen (z.B. de Bruin et al., 2015; Lehtonen et al., 2018) und potenzielle Vorteile im Alltag marginal sind (Grundy, 2020). Diese Studie untersuchte, ob Mehrsprachigkeit (d.h. das Sprechen von mehr als zwei Sprachen) die Leistung des Arbeitsgedächtnisses (AG) beeinflusst (Bialystok et al., 2012; Diamond, 2010). Zudem wurde geprüft, ob die Formalität des Spracherwerbs (formell vs. informell) das AG unterschiedlich beeinflusst.

Methoden. $N = 116$ gesunde Erwachsene (Alter: $M = 24.09$, $SD = 5.26$, Spannweite: 18.2–46.5 Jahre, 80 Frauen) wurden sechs AG-Aufgaben vorgelegt und der *Mehrsprachigkeitsgrad* mithilfe der Entropiegleichung nach Shannon (2001) berechnet. Die Formalität des Spracherwerbs ($FoLA_{ij}$) wurde anhand eines Fragebogens quantifiziert, wobei die Sprachkompetenz mit $FoLA_{ij}$ gewichtet und alternativ die Sprachen nach einem Grenzwert gefiltert wurden. Die AG-Komponentenstruktur wurde mittels Hauptkomponentenanalyse (PCA) und Parallelanalyse (PA) untersucht. Zusammenhänge zwischen Mehrsprachigkeitsgrad und AG-Leistung wurden mit linearen Regressionsmodellen (LRMs), kontrolliert für Alter, Bildung und Geschlecht, analysiert.

Ergebnisse. Die PCA/PA identifizierte eine numerisch-verbale AG-Komponente, die 47.11% der Varianz erklärt. Es gab keinen signifikanten Zusammenhang zwischen dem allgemeinen Grad der Mehrsprachigkeit und der numerisch-verbalen AG-Leistung, ebenso wenig zwischen den Mehrsprachigkeitsgraden, die mittels den $FoLA_{ij}$ gewichteten Sprachkompetenzen berechnet worden sind, und dem AG. Allerdings traten bei Anwendung von $FoLA_{ij}$-basierten Filtern die mittleren Schwellenwerte (vier und fünf) als signifikante Unterscheidungspunkte zwischen formellem ($p < .05$) und informellem ($p < .01$) Spracherwerb hervor. Auffallend war, dass über die gesamte Skala hinweg ein klares Muster erkennbar war: Formeller Spracherwerb wirkte sich positiv auf die AG-Leistung aus, während informeller Erwerb den gegenteiligen Effekt hatte.

Schlussfolgerung. Diese Ergebnisse deuten darauf hin, dass Spracherwerb nicht immer EF verbessert; vielmehr können die Auswirkungen auf das AG variieren: formelle Kontexte sind potenziell förderlich und informelle Kontexte möglicherweise hinderlich. Dies unterstreicht die Notwendigkeit, den Erwerbsstil als Variable in zukünftiger Forschung zur Wirkung der Mehrsprachigkeit auf die kognitiven Funktionen zu berücksichtigen.

Abstract

Background and Purpose. The reported positive impact of bilingualism on executive functioning (EF) remains controversial (e.g., Antoniou 2019), with meta-analyses showing none to minor effects (e.g., de Bruin et al., 2015; Lehtonen et al., 2018) and limited real-life benefits (Grundy, 2020). The present study extended this research to investigate whether speaking more than two languages (i.e., multilingualism) was related to working memory (WM) performance (Bialystok et al., 2012; Diamond, 2010). Breaking new ground, it was further examined whether the formality of language acquisition (formal vs. informal) was related to WM performance differently.

Methods. $N = 116$ healthy adults (age $M = 24.09$, $SD = 5.26$, range: 18.2–46.5 years, 80 female) were assessed on six WM tasks, and their degree of multilingualism was calculated via Shannon's entropy equation (2001). The formality of language acquisition ($FoLA_{ij}$) was quantified using items in the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007). To distinguish between formally vs. informally acquired languages, a complementary (weighting $FoLA_{ij}$ with speaking proficiency) and segmented approach (filtering languages based on $FoLA_{ij}$) was utilized. WM component structure was analyzed based on a principal component analysis (PCA) and parallel analysis (PA). Relationships between the degree of multilingualism and WM performance were examined with linear regression models, controlling for age, education, and sex.

Results. The PCA/PA identified a primary numerical-verbal WM component explaining 47.11% of the variance. There was no significant association between the overall degree of multilingualism and numerical-verbal WM performance nor between the degrees of multilingualism computed with $FoLA_{ij}$-weighted proficiencies and WM. However, when $FoLA_{ij}$-based filters were applied, the mid-range thresholds of four and five emerged as significant discrimination points for the relationship between formal ($p < .05$) and informal ($p < .01$) language acquisition and WM. Notably, a clear pattern emerged across all thresholds: formal language acquisition was related positively to WM performance, while informal acquisition had the opposite effect.

Conclusion. These findings imply that language learning might not always enhance EF; instead, its effects on WM could vary, with formal contexts potentially benefiting and informal contexts hindering WM performance. This highlights the need to consider the level of acquisition formality as a critical variable in future research on multilingualism's relation to cognitive functioning.

Exploring Working Memory in a Highly Multilingual Sample: The Degree of Multilingualism, Formality of Language Acquisition and Individual Variability in Working Memory

A positive impact of learning a second language (SLA, 2L) on executive functions/functioning (EF) has long been debated (e.g., Antoniou, 2019). EFs are defined as high-level cognitive processes enabling individuals to regulate thoughts and actions during goal-directed behavior (Friedman & Miyake, 2017). They consist of three core components: *shifting* (the ability to switch tasks or mental sets), *updating* (monitoring and revising working memory), and *inhibition* (suppressing impulsive responses), as featured in the “unity and diversity” framework by Miyake and Friedman (2012).

Some researchers carefully support the idea of a bilingual advantage in EF (Adesope et al., 2010; Grundy & Timmer, 2017; Grundy, 2020; van den Noort et al., 2019) and others challenge it (Donnelly et al., 2019; Lehtonen et al., 2018; Paap et al., 2017; Paap & Greenberg, 2013; Paap et al., 2015). While 54.3% of studies before 2019 reported a bilingual advantage in EF (van den Noort et al., 2019), meta-analyses reported negligible to small effects after accounting for publication bias (de Bruin et al., 2015; Lehtonen et al., 2018).

Research on the relationship between language learning and EF has mostly focused on bilinguals. However, considering that the bilingual advantages reported typically translate into just 30–100ms faster reaction times or 1–3% higher accuracy in EF tasks (Grundy, 2020), it raises the question of whether this qualifies as a truly significant benefit. It may, therefore, be valuable to assess individuals who speak more languages to determine if a greater impact is observable, which has previously been hypothesized (Bialystok et al., 2012; Diamond, 2010), though research on this topic remains relatively sparse. Kroll et al. (2012) further argue that “bilinguals, impressive as they are, are not special” (p. 254). For instance, enhanced working memory (WM) and attentional control have also been observed in musicians (D’Souza et al., 2018), gamers (Valls-Serrano et al., 2022), and athletes (Vaughan & Laborde, 2021). From a demographic perspective, Kroll et al.’s claim is supported by the fact that 73.3% of Europeans aged 25–34 and 55.1% of those aged 55–64 speak at least one foreign language (Eurostat, 2019), making bilingualism the norm rather than the exception.

In consequence, the current study sought to explore the relationship between the degree of multilingualism and WM, an important component of EF (Diamond, 2013), within a highly multilingual sample. Moreover, it was assessed whether and how the level of formality of language acquisition is related to WM performance. To provide some context, research on EF advantages in bilinguals will be reviewed before examining the question in multilinguals.

Bilingualism and Executive Functions

The Bilingual Advantage and Inhibition Control

Traditionally, proponents of the “bilingual advantage” (Bialystok et al., 2004) have emphasized the role of inhibition in bilinguals, drawing heavily on Green’s (1998) Inhibitory Control Model (ICM). The ICM posits that bilinguals must constantly suppress non-target languages when using a specific language, as all languages are simultaneously activated in the brain (Abutalebi & Green, 2007; Kroll et al., 2015). This continuous suppression, often over many years, strengthens inhibitory control and helps bilinguals avoid interference from non-target languages, a skill that likely also transcends to other EFs and cognitive domains (Abutalebi et al., 2008; Kroll et al., 2008; Tao et al., 2021).

The Shift Toward Attentional Control

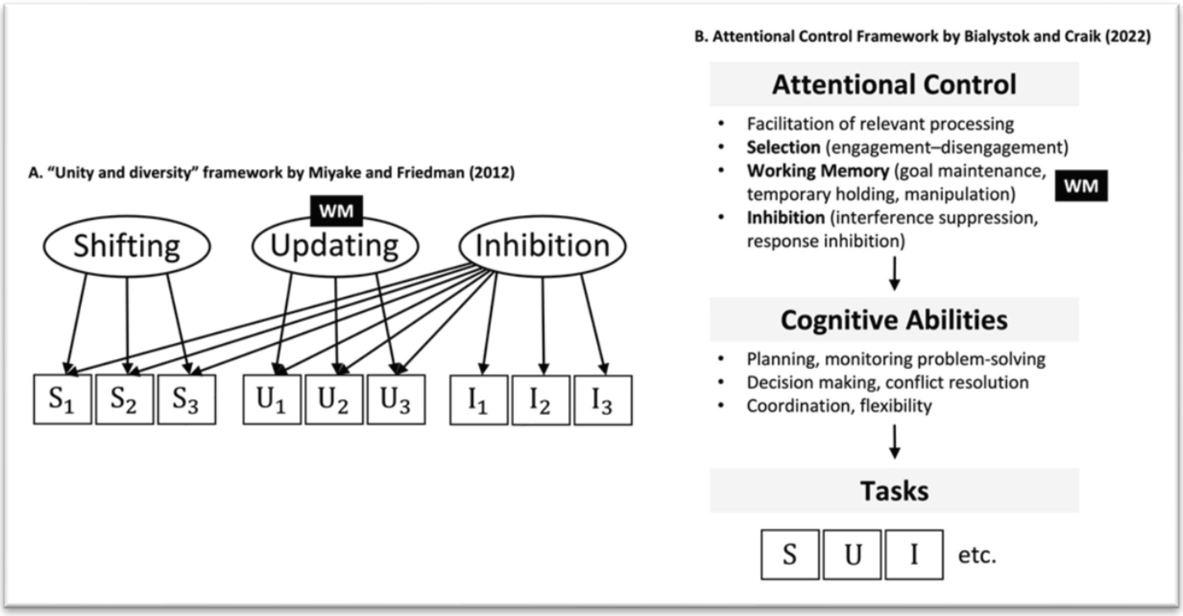
Despite the widely used ICM for explaining the bilingual advantage, the role of inhibition has faced scrutiny. For instance, Ware et al. (2020) demonstrated that bilingualism’s positive effects on EF are mainly seen in Attentional Network Tasks (ANT; de Souza Almeida et al., 2021) rather than in tasks measuring inhibition. EEG studies demonstrated larger P300 amplitudes and shorter latency during n-back tasks in bilinguals, providing empirical evidence for greater attentional resources (Comishen & Bialystok, 2021; Comishen et al., 2019). These findings align with Bialystok’s earlier reflections (2015, 2017), which suggested that attention, rather than inhibition, plays a crucial role in the bilingual advantage.

In light of this, the drawbacks of Miyake and Friedman’s (2012) “unity and diversity” framework for studying bilinguals have become increasingly evident, particularly due to its lack of focus on attention. To bridge this gap, Bialystok and Craik (2022) proposed a new framework that places attentional control at the top of cognitive abilities. This model merges processes like shifting, updating, inhibition, and WM into a cohesive, adaptable system. Unlike the modular design of Miyake and Friedman’s (2012) framework, where each process is largely isolated (apart from inhibition), Bialystok and Craik’s (2022) integrated hierarchical approach positions WM as an experience-driven resource within attentional control, influencing broader cognitive abilities and enabling efficient management of diverse tasks (see Figure 1). *Attentional control* is defined as a “repertoire of processing operations that specific tasks and higher-level cognitive functions can use to meet their various goals”

(Bialystok & Craik, 2022, p. 1252). It is closely related to the construct of *executive attention*; however, for simplicity, this study uses *attentional control* and *executive attention* interchangeably. To understand how bilingualism influences WM and relates to attentional control, a brief overview of the most widely accepted WM model is provided.

Figure 1

Comparison of the "Unity and Diversity" Framework and the Attentional Control Framework



Note. In the "unity and diversity" framework by Miyake and Friedman (2012), inhibition is considered a common EF across tasks, while each component (shifting, updating, and inhibition) shows task-dependency, with specific tasks linked to each process. In contrast, Bialystok and Craik's (2022) attentional control framework lacks task-specificity, viewing attentional processes as flexible and adaptable rather than tied to specific tasks. Note the different allocation of WM: In Miyake and Friedman's model, WM is mainly associated with the updating component, while in Bialystok and Craik's framework, WM is integrated within attentional control.

U = updating task; *I* = inhibition task; *S* = shifting task.

Bilingualism, Working Memory, and Attention

The Multicomponent Model of Working Memory

In the most general terms, working memory (WM) is a cognitive system that temporarily holds and processes a small set of information in an active and accessible state while executing complex tasks (Baddeley, 2010; Cowan, 2017). To date, numerous models have been proposed to understand the structure of WM with the multicomponent model of WM by Baddeley and Hitch (1974) being perhaps one of the most prominent ones in the WM literature (Adams et al., 2018). The multicomponent model of WM features a *central executive* that manages attention, along with three subsidiary components: *the visuo-spatial sketchpad* for visual sensory input, the *phonological loop* for verbal-acoustic content, and the *episodic buffer* that consciously integrates diverse information, creating a coherent "multidimensional code" accessible by all components of WM and long-term memory (LTM) (Baddeley, 2000, 2003, 2010, 2012; Baddeley & Hitch, 1974). This structure of WM has been validated through factor analyses, which identified separate numerical-verbal and visuospatial components (Waris et al., 2017).

In contrast to LTM, WM, as transient storage, is characterized by its constrained capacity to maintain information, especially if adequate strategies to combat memory trace decay are not utilized (Cowan, 1998). For temporal decay, individuals may engage in subvocal rehearsal, i.e., repeating verbal items by inner voice (Buchsbaum & D'Esposito, 2008), or apply attentional refreshing, i.e., quickly refocusing on single memories to sustain them (Barrouillet et al., 2004; Barrouillet & Camos, 2015).

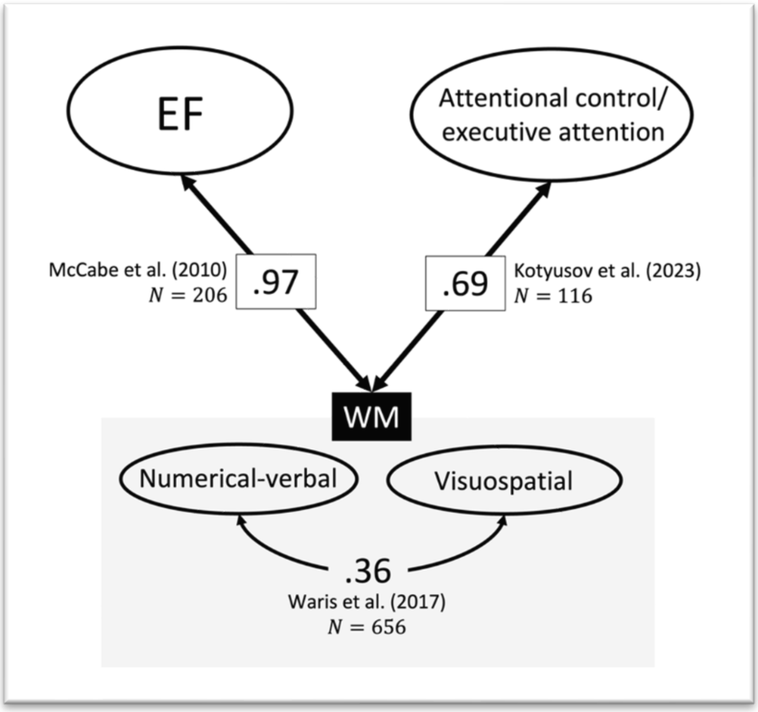
When discussing WM capabilities, two main distinctions are generally made: "simple" *WM performance*, measured by tasks like serial number or word recall, and *WM capacity* (WMC), which involves more complex tasks that require both storing and manipulating information (Barrett et al., 2004). This study used the term *WM performance* holistically, as a wide range of WM tasks was incorporated.

The Interplay Between WM and Attention in the Bilingual Brain

Research has demonstrated that WM is essential for SLA (Baddeley, 2015) and strongly tied to EF (Friedman & Miyake, 2017; Miyake & Friedman, 2012; Miyake et al., 2000) if not considered part of EF itself (Diamond, 2013). As expected, McCabe et al. (2010) found a strong correlation ($r = .97$) between WMC and EF, suggesting that these constructs are closely intertwined. Notably, they also observed that complex span tasks tap into a shared underlying mechanism of executive attention. Engle (2002) found that WMC involves not just storing information but also controlling attention. Furthermore, Kane et al. (2007) demonstrated that while n-back tasks are frequently used to measure WMC, they showed stronger convergent validity with attentional control tasks, especially in older age (Gajewski et al., 2018). These insights reinforce the idea that both WM and EF rely heavily on the same cognitive process of attention, as depicted in Figure 2.

Figure 2

Related Constructs of WM, EF, and Attention



Note. Correlations were not weighted for sample size; this figure presents simplified relations.

Research suggests that bilinguals often outperform monolinguals in various WM tasks (Bialystok, 2017). For example, Blom et al. (2014) discovered that bilingual children and young adults showed superior performance in backward WM tasks, such as the digit span task, but no advantage in forward tasks. Antón et al. (2019) echoed these findings, adding that bilinguals only differed from monolinguals in visuospatial and numerical-verbal WM when tested on backward WM tasks. Furthermore, Janus and Bialystok (2018) noted greater accuracy but slower responses in bilingual children on n-back tasks compared to monolinguals. These results underscore a bilingual advantage, particularly in executive attention, since backward WM tasks involve reordering information, requiring more cognitive control and relying more heavily on attentional processes (Kotyusov et al., 2023).

Calvo et al. (2016) argued that the bilingual advantage appears not to be equally evident across all WM components. Indeed, some studies (Bialystok et al., 2014; Luo et al., 2013; Morales et al., 2013; Morrison et al., 2019) have found that bilinguals show an advantage over monolinguals in *visuospatial* WM tasks, but monolinguals frequently outperform bilinguals in numerical-verbal WM tasks. This difference could be due to the higher lexical demands of verbal tasks, which may create an additional cognitive load for bilinguals managing multiple language representations, favoring monolinguals (Bialystok et al., 2014). However, other studies have reported a positive impact of bilingualism on both numerical-verbal and visuospatial WM components (Blom et al., 2014; Warmington et al., 2019). Conversely, several studies have found no consistent evidence for such effects at all (Bonifacci et al., 2011; Engel De Abreu, 2011; Lehtonen et al., 2018; Lukasik et al., 2018; Namazi & Thordardottir, 2010; Ratiu & Azuma, 2015; Yang, 2017), reflecting the mixed results seen in other EFs.

Meta-analyses on the relationship between bilingualism and WMC reported significant small to medium effect sizes of $r = .20$ (Grundy & Timmer, 2017) and $r \approx .23/\ g = .48$ (Adesope et al., 2010). As mentioned in the beginning, this implies that while there is probably an effect, its presence has not been consistently observed across studies or may be relatively weak in bilinguals.

What about Multilinguals?

The question at hand of whether and how speaking multiple languages affects WM differently compared to mono- and bilingualism has sparked controversy amongst researchers. At this point, two lines of argumentation have emerged: one suggests that the cumulative cognitive demands of juggling multiple languages further enhance EF (Bialystok et al., 2012; Diamond, 2010). Conversely, managing more than two languages could also lead to even greater interference and competition between language representations, potentially weakening EF. In the following, a summary of findings on multilingualism and WM from infancy to old age is offered.

Brito et al. (2014) demonstrated that bilingual infants exhibited superior WM performance compared to monolinguals, whereas trilinguals performed similarly to monolinguals. However, in children aged five to eight, Poarch and van Hell (2012) reported that trilinguals and simultaneous bilinguals outperformed second-language learners on the

ANT task, though no significant difference in attentional control was observed between tri- and bilinguals. In university students, Cockcroft et al. (2019) showed a multilingual advantage across all WM components, with the strongest effects found in processing-heavy tasks like the backward digit span task and spatial recall tasks. Additionally, Espi-Sanchis and Cockcroft (2022) noted that balanced language proficiency (i.e., being equally adept in multiple languages) in young adults significantly improved verbal, but not visuospatial, WM. Fyndanis et al. (2023) found that, among academics aged 54 to 70, the number of languages spoken significantly predicted verbal WMC, as measured by the digit span task, although visuospatial WM was not assessed.

As put forth by Perquin et al. (2013), lifelong multilingualism, especially when initiated early and practiced throughout life, can offer protection against cognitive impairments. Stronger effects were observed as the number of languages increased, particularly from bilingualism to trilingualism. For instance, Chertkow et al. (2010) reported that speaking three or more languages delayed Alzheimer's disease onset by five to seven years compared to bilinguals and monolinguals. Kavé et al. (2008) found that older trilinguals and quadrilinguals outperformed bilinguals on geriatric cognitive tests like the Mini-Mental State Examination (MMSE; Folstein et al., 1975). Similarly, Achaa-Amankwaa et al. (2023) established that multilinguals older than 60 years excelled in attentional control-demanding tasks like the backward digit span task, though the effects were smaller in other WM tasks, indicating a task-specific benefit.

Given these studies explicitly examined multilinguals, one would expect the effects of multilingualism to be most pronounced in the *numerical-verbal* WM component, particularly in tasks demanding high levels of attentional control and therefore the engagement of the central executive within the multicomponent model of WM. These effects should increase by the number of languages spoken (Fyndanis et al., 2023), especially in cases where individuals exhibit balanced proficiency (Espi-Sanchis & Cockcroft, 2022).

Capturing the Multilingual Experience: Language Entropy

Even within an apparently homogenous group of bilinguals, the effects on WM vary. For instance, Yang (2017) found that intermediate bilinguals performed significantly better on the digit span task than high-proficiency bilinguals, whose results were similar to those of a near-monolingual group. This indicates that the level of L2 proficiency varies with WM performance effect sizes between monolinguals and bilinguals and points out the importance of assessing bi- and multilingualism in a nuanced, continuous manner, as variations in language proficiency levels can significantly influence WM performance.

Luk and Bialystok (2013) argued for conceptualizing bilingualism as a multidimensional phenomenon, as they found a significant interaction between language proficiency and use, highlighting the need to assess language experience continuously in terms of proficiency and usage rather than as a simple categorical status. Hence, the present study implemented a continuous and more practical method for assessing language proficiency in multilingual settings. Self-reported data were used, the degree of multilingualism was modeled in a unique two-dimensional way (see Methods).

This approach included applying the concept of *entropy*, a measure from information theory to the study of languages. Entropy continuously quantifies the average uncertainty or degree of unpredictability in a random variable's outcome (Shannon, 2001). Within multilingualism research, entropy was initially introduced as a measure of language context variability (Gullifer et al., 2021; Gullifer & Titone, 2020, 2021a, 2021b; Tiv et al., 2020). Another use case is found in Kepinska, Dalboni Da Rocha, et al.'s (2023) study, which employs entropy to assess the diversity and duration of exposure of participants' language experiences. In this study, entropy is interpreted as an indicator of whether a participant maintains equal competence across several languages or concentrates proficiency in a few languages (see Table 1 and Methods section for further explanation).

Table 1

Shannon's Entropy H_j Formula and Applications in Language Studies

$$\left(p_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}}\right) \qquad H_j = -\sum_{i=1}^{n} p_{ij}\, log_2\left(p_{ij}\right)$$

	Gullifer et al.	Kepinska, Dalboni Da Rocha, et al. (2023)	Present study
Calculation of p_{ij}	$p_{ij} = \frac{\sum_c u_{ic}}{\sum_c \sum_i u_{ic}}$	$x_{ij} = -[\log(AoA_{ij} + 1) - max(AoA_{ij})] + 1$	$x_{ij} = a_{ij}$
Focal point	Diversity of language use in social contexts	Diversity of language AoA and duration of exposure	Spread and level of language proficiency
High entropy interpretation	Person switches between many languages in the same social context (compartmentalized)	Person exposed to multiple languages simultaneously early in life	Person's proficiency is evenly distributed across multiple languages
Low entropy interpretation	Person uses the same languages across different social contexts (integrated)	Person exposed to few languages early in life, learning additional languages successively later in life	Person's proficiency is concentrated in fewer languages

Note. For accurate entropy calculation, p_i values must represent probabilities, meaning an input variable x_i may require normalization beforehand if it is not already an empirical probability. i = language index; j = participant index; n = total number of languages subject is exposed to; u_{ic} = usage of language i in a specific social context c; AoA_i = age of acquisition of language i; a_i = self-reported proficiency in language i. The equation for Gullifer et al.'s entropy is based on the general framework as described in Gullifer et al.'s work. Entropy (H_j) is calculated for each participant j individually.

A Question of Formality?

On a more exploratory note, this study also examined the potential effects of formal versus informal language acquisition on WM performance. Generally, *formal learning* takes place in structured educational settings like schools or universities, with prescribed frameworks and defined outcomes (Hager, 2012), while *informal learning* occurs organically through daily experiences, without formal frameworks, teacher-led instruction, grades, or qualifications (McGivney, 2001). However, these labels should not be seen as rigid categories (Johnson & Majewska, 2024), as learning boundaries are becoming increasingly blurred, particularly with advancements in digital technology, such as social media or (language) learning apps (Inayati et al., 2024).

Regarding WM, Souza-Talarico et al. (2007) reported that for individuals over 60 years old, the number of years in formal schooling was positively associated with performance on the backward digit span task. Similarly, Kosmidis et al. (2011) found that educated literates performed better than self-educated literates (i.e., individuals who learnt writing and reading outside of the formal education system) on the backward digit span task. In younger populations, Roberts et al. (2015) discovered that classroom exposure, rather than age, was a reliable predictor of WM performance in 6- to 7-year-olds. In their 2023 study, Davidson et al. found that children aged 4 to 7 experienced significant improvement in WM performance during their first year of formal schooling. Finch (2019) noted stronger WM performance gains throughout the school year compared to the summer break. Collectively, these studies underscore the relevance of formal education in improving WM performance, especially when contrasted with informal or less structured learning contexts.

With respect to language learning, the role of formality has been acknowledged (De Wilde et al., 2022; Dressman & Sadler, 2020; Krashen, 1976); however, there appears to be a lack of studies examining the association between language learning formality and WM performance.

The Present Study

This study explored the effects of multilingualism on WM performance and, for the first time, examined whether proficiency in formally versus informally acquired languages is linked to WM performance differently. To ensure the presence of the common WM structure in the sample, a principal component analysis (PCA) was conducted on various WM tasks, anticipating a two-component structure (Waris et al., 2017). The degree of multilingualism was measured continuously, as recommended by Luk and Bialystok (2013), using Shannon's (2001) entropy equation applied to language proficiency data.

Given the assumption that speaking more than two languages enhances EF efficiency (Bialystok et al., 2012; Diamond, 2010), a positive association was predicted specifically between the degree of multilingualism and the numerical-verbal WM component. Among bilinguals, studies reporting a bilingual advantage have identified various combinations of cognitive benefits, and showed effects to be more pronounced in the visuospatial WM component, or both, numerical-verbal and visuospatial WM (Bialystok et al., 2014; Blom et al., 2014; Luo et al., 2013; Morales et al., 2013; Morrison et al., 2019; Warmington et al., 2019). For multilinguals, however, research observed cognitive advantages more frequently in the numerical-verbal WM component (Cockcroft et al., 2019; Espi-Sanchis & Cockcroft, 2022; Fyndanis et al., 2023), which directed the predictions toward numerical-verbal WM performance (see Table 2). Although the primary focus was on the numerical-verbal component of WM, a variety of tasks targeting other WM components were included (see Table 3) to assess whether the two-component WM structure could also be reproduced in this study's sample.

Table 2

Selected Studies Reporting WM Benefits in Bilinguals and Multilinguals

	Study	Numerical-verbal WM task	Visuospatial WM task
Bilinguals	Bialystok et al. (2014)	○	●
	Blom et al. (2014)	●	●
	Luo et al. (2013)	○	●
	Morales et al. (2013)		●
	Morrison et al. (2019)		●
	Warmington et al. (2019)	●	●
Multilinguals	Cockcroft et al. (2019)	●	●
	Espi-Sanchis and Cockcroft (2022)	●	○
	Fyndanis et al. (2023)	●	

Note. The table was adapted and expanded from Liu and Liu (2021), following their selection criteria with modifications to include multilingual participants (continuous scale accepted) and focus on studies showing significant effects in one or both task types, without requiring detailed participant descriptions. Studies with inconclusive findings are excluded. This table offers a simplified overview, omitting details on sample sizes, populations, ages, control variables, and specific task types. For a comprehensive review, see Grundy and Timmer (2017). Filled circles (●) indicate studies showing significantly better performance in WM tasks among bilinguals or multilinguals compared to monolinguals/bilinguals. Empty circles (○) indicate no observed improvement in performance, and blank cells indicate that no tasks of that type were administered.

Furthermore, a novel distinction between formally and informally acquired languages was introduced, investigating whether this difference was also reflected in varying effects on WM. The research questions and hypotheses were as follows:

RQ 1: What is the underlying structure of different WM tasks in a highly multilingual sample?

H1: Different WM tasks correlate at least moderately ($r \geq 0.30$; Cohen, 1988) with each other.

H2: The different WM tasks can be clustered into at least two compound measures, i.e., components of WM: numerical-verbal and visuospatial.

RQ 2: How are the degree of multilingualism and WM performance related to each other?

H3: The degree of multilingualism is a stronger predictor of performance in numerical-verbal WM than in visuospatial WM.

H4: The degree of multilingualism in formally acquired languages is associated with WM performance differently than in informally acquired languages.

Methods

Participants

The total sample consisted of $N = 116$ healthy adults ($M = 24.09$, $SD = 5.26$, range 18.2–46.5 years), including 80 females and 36 men with 15.91 years of education on average. This study used the data of individuals which have been already recruited by Rampinini et al. (2024), yet excluding dyslexics ($n = 29$) and hyperpolyglots ($n = 7$). Subjects were recruited from the Geneva region via promotional flyers and online advertisements. Individuals with professional qualifications in either music or simultaneous interpreting were excluded. Most were native French speakers, but those proficient in more than six languages were eligible if they were non-native French speakers with advanced French skills. All participants who met the inclusion criteria signed informed consent forms authorizing all experimental procedures and the future reuse of data within the open science framework.

Procedure

To assess WM performance, multiple WM tasks were deployed. All behavioral tasks were run in the cloud-based research platform Gorilla Experiment builder (www.gorilla.sc). Instructions were provided in French, displayed on the computer screen, and read out aloud by an AI-generated voice ("Renee France", speed = –1, www.naturalreaders.com). Participants had no option of skipping the instructions. If protocol accommodated it, sample trials were presented. The digits in the digit span task were presented by the same AI voice as the instructions. The non-words in the non-word repetition task were voiced by a pre-recorded female voice. The spoonerisms were read by a pre-recorded male voice. If not indicated differently, tasks were adapted for computer usage, scoring simultaneously and verified based on a recording of the participants' answers.

Measures and Materials

Working Memory

The *Corsi Block-Tapping Test* (CBT; Corsi, 1972; Milner, 1971) is typically used to evaluate visuospatial working memory (VSWM). In its traditional form (Arce & McMullen, 2021), participants observe a flat surface on which nine cubes are randomly positioned. The examiner taps these cubes in a specific order, which participants must then replicate. Participants repeat the tapping sequence as it was presented for the forward task and in reverse order for the backward task. The length of the sequence is increasing up to the point where the participant fails to reproduce a sequence. This is referred to as the "Corsi span" and reflects the capacity of VSWM. In this study, a digital 2-dimensional adaption of the CBT was deployed (layout based on Kessels et al., 2000) consisting of a computer display featuring nine black squares positioned on a white background. Sequences were highlighted in yellow, and selections made by mouse-click were indicated in green. As Brunetti et al. (2014) and Siddi et al. (2020) reported, the traditional and digital versions of CBT produce comparable results, suggesting that the digital CBT is a valid and reliable measurement of VSWM.

The *digit span task* serves as a standard measure of auditory working memory (AWM; Richardson, 2007) and is a component of the Wechsler Adult Intelligence Scale-Fourth Edition (Wechsler, 2008). In this task, participants hear sequences of digits recited by the examiner. They are then asked to repeat each sequence as presented (forward digit span task) or in reverse (backward digit span task). After correct reproduction, the sequence of digits increases each trial until the participant fails to recall the complete sequence. The maximum number of digits successfully repeated is called the "digit span". For this work, the digit presentation was conducted through headphones and participants entered the sequence via the upper numeric keyboard.

During the *spoonerism task*, participants are presented with audio recordings of word pairs. They are required to exchange the first sounds (phonemes) of each word and articulate the emerging nonwords in their initial order (Szenkovits & Ramus, 2005); for example, "pipe—robe" becomes "ripe—pobe". Spoonerisms allow assessing an individual's phonological awareness (PA), i.e., the capability to detect and alter the sound structure of spoken language, which is closely linked to literacy (Anthony & Francis, 2005). They are essentially speech errors (Harley, 2006), often utilized in the context of examining dyslexics (e.g., Ramus, 2003; Varvara, 2014); however, given their association with WM (Knoop-van

Campen et al., 2018; Richardson et al., 2011), spoonerisms can also be employed as a suitable measure for PWM. In the present study, stimuli from Szenkovits and Ramus (2005) were used in a computer-based interface where a natural reader provided the instructions and word pairs. Answers were recorded via microphone and scored by the examiner.

The *Advanced Measures of Music Audiation Test* (AMMA; Gordon, 1989) is a standardized test originally intended to investigate musical aptitude. It comprises of 15 pairs of short melodies, each altered either in melody or rhythm. Participants are tasked with judging whether alterations are melodic, rhythmic, or absent. Tonal and rhythmic accuracy scores are calculated and aggregated into a total score. Despite being initially designed to evaluate audiation, a study by Platz et al. (2022) revealed that just five out of the 30 items used in AMMA yield a model fit that constitutes invariant measurement when compared to the *Measures of Musical Talents Test* (MMTT; Seashore, 1919)—a historic test designed to measure tonal memory. Furthermore, the AMMA and MMTT showed a high correlation ($Mdn_{\tau} = .5,\ 95\%\ CI\ [0.42, 0.70]$; Platz et al., 2022). As such, the AMMA scores appear to be confounded with WM, which is why this study used it as a WM measure. The AMMA was modified for computer-based presentation, yet its contents remained unchanged.

The *non-word repetition task* (NWRT; Coady & Evans, 2008) is widely used to examine phonological working memory (PWM; e.g., Gathercole et al., 1994). In this task, participants are exposed to lists of pseudowords—made-up words similar to those in their native language—and must repeat them back immediately. In the current study, following Majerus & Van der Linden (2003), participants went through 24 computer-based trials, and a raw accuracy score was calculated based on how many pseudowords were repeated correctly, irrespective of their serial position.

The *California Verbal Learning Test* (CVLT; Delis et al., 1987) is a standardized test battery for examining verbal learning and verbal working memory (VWM; Delis et al., 1988; Lacritz & Cullum, 1998). In this test, participants memorize a shopping list read aloud by the examiner containing 16 words. Initially, participants attempt to remember as many words as possible over five trials (immediate recall). Subsequently, a new shopping list is introduced once (interference), which they need to recall before returning to remember the initial list, excluding any words from the interference list (short-delay recall). After a 20-minute delay, participants must recall the original list (long-delay free recall) and categorize the words into four provided semantic groups (long-delay cued recall). Finally, participants are asked to remember words from the original shopping list based on a combined pool of words from both the original and interference list orally presented to them (recognition trial). To assess VWM, an aggregated score from the immediate and short-delay recall was used, given that the 20-minute long-delay recall already appears to tap into LTM consolidation processes (Ladowsky-Brooks, 2016). A computer-based version was programmed, following the French version of CVLT-II adapted by Deweer et al. (2008).

Degree of Multilingualism

As mentioned in the introduction, a novel approach was employed to measure the degree of multilingualism. Specifically, the *entropy* R-package (v1.3.1; Hausser & Strimmer, 2021) was used. To estimate entropy, the *entropy* function first normalizes each individual's proficiency score for a given language by dividing it by the total proficiency across all languages for that individual (maximum likelihood estimation). This produces the necessary empirical probabilities, which are then put into Shannon's (2001) entropy formula:

$$p_{ij} = \frac{a_{ij}}{\sum_{i=1}^{n} a_{ij}} \quad \rightarrow \quad H_j = -\sum_{i=1}^{n} p_{ij}\, log_2\left(p_{ij}\right)$$

Here n represents the total number of languages a participant j is exposed to, i is the language index, a_i reflects the self-reported speaking proficiency in language i, p_i is the relative proficiency of the i-th language across the total sum of language proficiencies, and H_j is the entropy of participant j, representing the spread of language proficiency across different languages. Therefore, the final equation could be conceptualized as the *degree of multilingualism* calculated from the distribution of proficiency across multiple languages rather than from generic probabilities. It measured how diverse and balanced (*high* degree of multilingualism) vs. how specialized and concentrated (*low* degree of multilingualism) language proficiency in participants was; in short, it captured two dimensions: the spread and level of language proficiency.

Formality of Language Acquisition

A new variable, *formality of language acquisition* ($FoLA$) was created. The rationale was to establish a variable that indicates the degree of formality in language learning, from casual interactions with family and friends (informal) to structured educational settings like classroom learning or instructional teaching (formal). For this purpose, four items in the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007) were selected: contribution of family, friends, life, and education on language learning. The education scale was inverted to ensure that the final variable was unidirectional. The four items were summed per language i and participant j then divided by the total number of valid responses, with missing data points excluded:

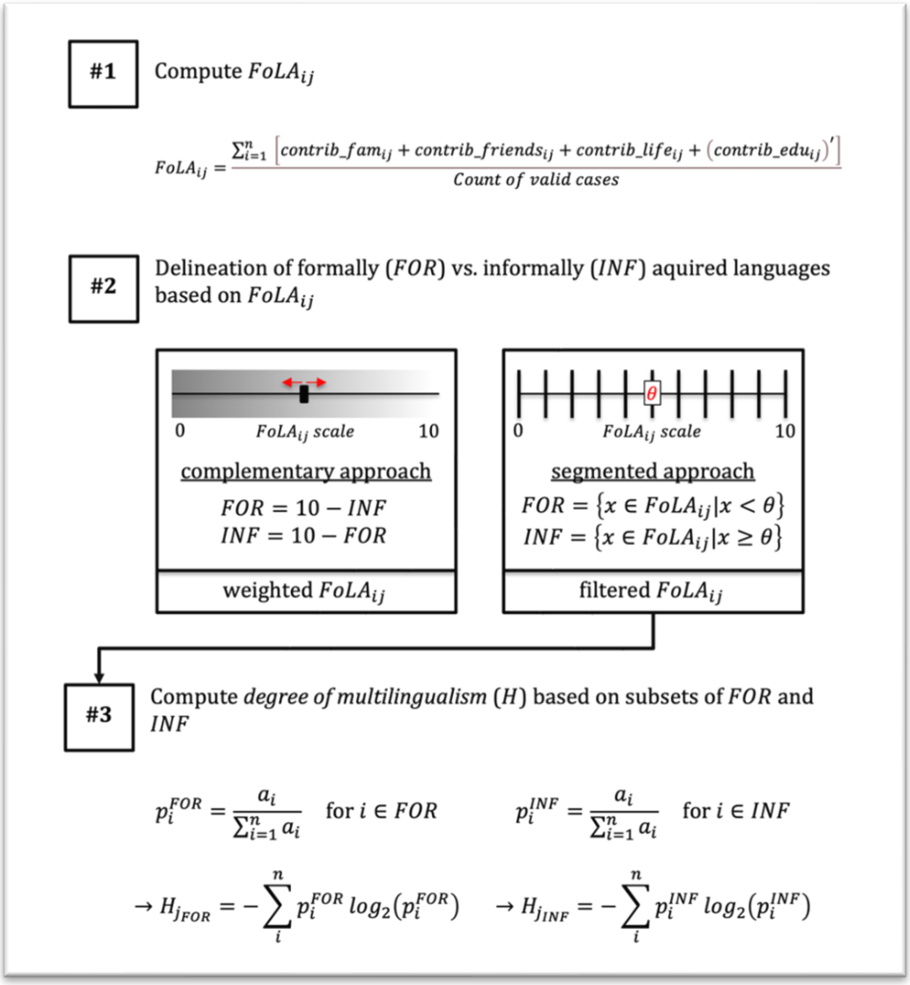
$$FoLA_{ij} = \frac{\sum_{i=1}^{n} \left[contrib\_fam_{ij} + contrib\_friends_{ij} + contrib\_life_{ij} + \left(contrib\_edu_{ij}\right)'\right]}{Count\ of\ valid\ cases}$$

$FoLA_{ij}$ has a range of 0–10, whereas *low* scores indicate a high intensity of formal language instruction, while *higher* scores imply the respective language was learnt in an informal setting.

$FoLA_{ij}$ was then used to compute two versions of the degree of multilingualism: one based on formally learnt languages (FOR) and the other based on informally learnt languages (INF). As such, each participant was assigned two variables for the degree of multilingualism. The concept of language acquisition formality has not been formalized in this way before. Hence, the challenge was determining a numeric delineation (cut-off, threshold, continuum, gradient, etc.) by which a given $FoLA_{ij}$ score reflects (more) formal or informal acquisition. A data-driven approach was chosen, comprising of two strategies: first, the degree of multilingualism was computed based on weighted $FoLA_{ij}$ scores (*complementary approach*), i.e., each participant's $FoLA_{ij}$ score was multiplied by the proficiency in the respective language. However, since the most dominant language is usually acquired informally, one can expect that the distribution of $FoLA_{ij}$ scores will be skewed and will potentially distort any relationship between the degree of multilingualism and WM performance. Thus, manual thresholds (θ) were additionally set along the $FoLA_{ij}$ scale (scale points 2–9) for filtering formally vs. informally acquired languages (*segmented approach*). The workflow is depicted in Figure 3.

Figure 3

Steps for Creating FOR and INF Subsets



Note. x = value on $FoLA_{ij}$ scale; θ = set threshold; i = language index; j = participant index; a_i = self-reported proficiency in language i; p_i = relative proficiency of the i-th language across the total sum of language proficiencies. Step #3 is depicted only for $FoLA_{ij}$-filtered proficiencies, as $FoLA_{ij}$-weighted proficiencies included additional intermediate steps.

In conclusion, this study examined the constructs of WM, degree of multilingualism, and $FoLA_{ij}$ based on using seven measures/materials as summarized in Table 3, and an overview of all variable codes is included in the Appendix.

Table 3

Overview of Constructs and Associated Measurements/Materials Analyzed in this Study

Construct	WM Component	Measure/Material
Working Memory (WM)	Auditory	Digit Span Forward and Backward
	Auditory	Non-Word Repetition (NWRT)
	Verbal	California Verbal Learning Test (short) (CVLT)
	Phonological	Spoonerisms
	Tonal	Advanced Measures of Music Audiation Test (AMMA)
	Visuospatial	Corsi Block-Tapping Test (CBT) Forward and Backward
Entropy		Language Experience and Proficiency Questionnaire (LEAP-Q) $FoLA_{ij}$ score as filter or weight for entropy calculation
Formality of Language Acquisition ($FoLA_{ij}$)		Language Experience and Proficiency Questionnaire (LEAP-Q)

Statistical Analyses

All analyses were performed in R statistical software (v4.1.1; R Core Team). The R packages *misty* (v0.6.5; Yanagida, 2024), *readr* (v2.1.5; Wickham, Hester & Bryan, 2024), *purrr* (v1.0.2; Wickham & Henry, 2023), *tidyr* (v1.3.1; Wickham, Vaughan & Girlich, 2024), and *dplyr* (v1.1.4; Wickham, François, et al., 2023) were used for data wrangling, modification, and cleaning. Multivariate normality was assessed with the *MVN* package (v5.9; Korkmaz et al., 2014). Correlation analyses were computed by applying the *corrplot* package (v0.92; Wei & Simko, 2021) and PCA by the *psych* package (v2.4.6; Revelle, 2024). Following Kepinska, Dalboni Da Rocha, et al. (2023), entropy was calculated using the *entropy* package (v1.3.1; Hausser & Strimmer, 2021). Statistical modeling was conducted based on the *mgcv* (v1.9-1; Wood, 2011) and *lme4* (v1.1-35.3; Bates et al., 2015) packages. For subsequent analyses, a significance level of 5% was chosen. Thus, results with p-values of .05 or below were deemed statistically significant.

First, the data files obtained from Rampinini et al. (2024) were merged into one data frame containing only variables relevant to this study. There was one instance of missing data, for which the mean was imputed for the NWRT. Second, normality was assessed using the Mardia test (Mardia, 1970), the Anderson-Darling test (Stephens, 1974), and the Shapiro-Wilk test (Shapiro & Wilk, 1965). Apart from the AMMA variables and non-word repetition accuracy, all other variables did not follow a normal distribution. Third, outliers were checked for using the Median Absolute Deviation method (MAD; Wilcox, 2022) with ±3 MADs. This method is supposed to be less sensitive to extreme values and more appropriate for skewed or nonnormal data (Leys et al., 2013). Only variables with extreme outliers that profoundly skewed the distribution were checked. In the present case, the variables of cvlt-long_RT, spoon_acc and spoon_RT_manual were *winsorized*, i.e., their values were adjusted to the nearest value within the 5th and 95th percentiles (Riffenburgh, 2012). This approach ensured that outliers were not entirely excluded, which is essential to maintaining an adequate sample size. As pointed out by Zhou et al. (2016), rigorous data trimming could obscure the cognitive advantages being measured in multilinguals. Lastly, two variables (reaction time for CVLT and Spoonerisms) were inverted, and all variables were z-standardized to ensure each variable contributed equally to the analyses regardless of its scale.

To assess the underlying structure of WM in the current sample (H1 and H2), a Spearman correlation matrix was computed, including all WM measures. This also allowed for the exclusion of potential singularities (i.e., cases of extreme multicollinearity), non-significant variables, and redundant variables from further analyses. Also, except for spoon_RT_manual, all WM measures were found to be linearly associated with each other. Next, a principal component analysis (PCA; Pearson, 1901; Hotelling, 1933) was performed to detect the principal components capturing the highest amount of variance in the data. Before PCA computation, a parallel analysis (PA; Horn, 1965) with 1,000 iterations was simulated to identify the optimal number of components to extract. The components whose actual eigenvalues were greater than the eigenvalues derived from the resampled data were retained. To facilitate subsequent analyses, WM performance was then described as one or more component scores. PCA was chosen because, unlike maximum likelihood estimators used in EFA or CFA, it does not have distributional assumptions, making it more suitable in cases of nonnormality (Fabrigar et al., 1999). Furthermore, the PA is more commonly applied and evaluated in the context of PCA, supporting its validity (Crawford et al., 2010; Steger, 2006). With no consensus on the ideal sample size for dimensionality reduction (Kyriazos, 2018), the Kaiser-Meyer-Olkin test (KMO) was performed to determine the sample's suitability as recommended by Field (2012). An overall MSA value of 0.74 indicated "good" adequacy and large enough correlations between the variables.

To examine the relationship between the degree of multilingualism and WM performance (H3 and H4), linear regression models (LRMs) and generalized additive models

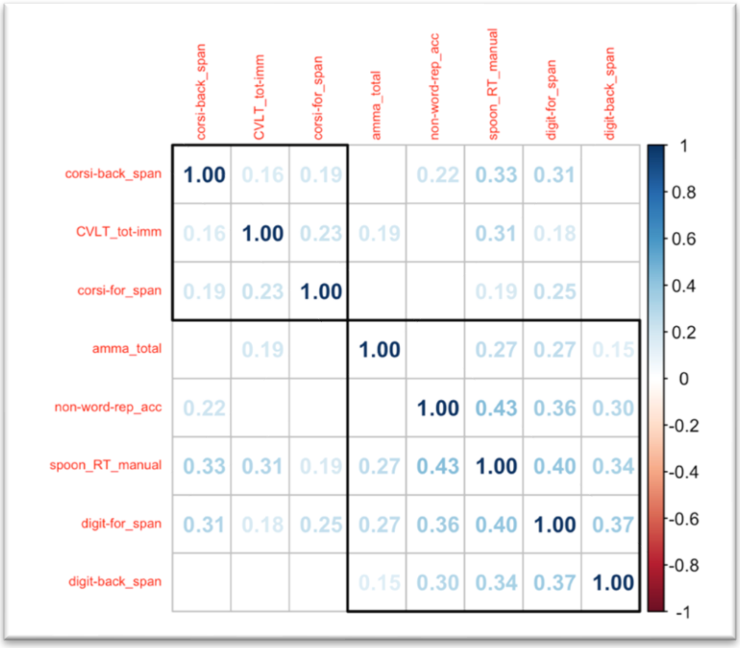
(GAMs; Baayen & Linke, 2020; Hastie & Tibshirani, 1986) were computed with global entropies, $FoLA_{ij}$-weighted, and $FoLA_{ij}$-filtered entropies as single predictors, controlling for age, sex, and education. WM performance (now consolidated in component scores) served as the response variable. GAMs were included because they expand on traditional linear models by permitting each predictor to influence the outcome through a smooth function, thus enabling a more flexible fit in case of non-linearity. If the GAMs provided a poorer fit than the LRMs, the results were interpreted, and effect sizes were calculated exclusively from the LRMs.

Results

The Spearman correlation matrix revealed high correlations ($r \geq .82$) between the variables CVLT_tot-imm, CVLT_short, amma_total, amma_tonal, and amma_rhythm, indicating singularity. As such, the correlation matrix was non-invertible for component analysis, and high redundancy was apparent. Based on the correlation coefficients and theoretical implications, only CVLT_tot-imm and amma_total remained in subsequent analyses. The variables cvlt_long-RT and spoon_acc showed only a few significant correlations with other variables, so they were excluded. Using the remaining WM measures, another Spearman correlation matrix was computed, as presented in Figure 4. The CBT backward exhibited the overall highest total correlation and most widespread correlations with other WM measures, whereas the backward digit span task was linked only to its forward counterpart. Of the 28 correlations tested, 20 were significant (71.43%), and nine of these significant correlations (45%) yielded $r \geq 0.30$.

Figure 4

Spearman Correlation Matrix of Final WM Task Selection

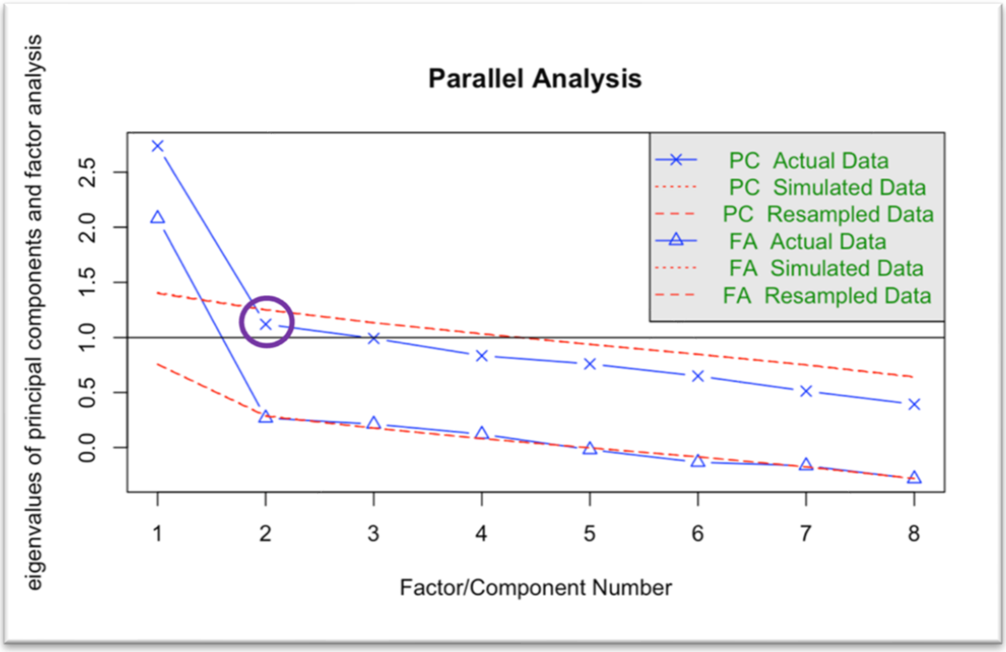


Note. Spearman correlations sorted from highest to lowest. All correlations depicted are significant ($p \leq 0.05$).

As previously stated, the final selection of WM tasks (see Figure 4) yielded a sufficient overall MSA value ($> \mathbf{0.50}$), indicating suitability for PCA. Based on the PA, one component could be extracted with confidence. However, a second component could not be derived, as its PC eigenvalue fell *below* the eigenvalue obtained from the PC resampled data (see Figure 5).

Figure 5

Results of Parallel Analysis



Note. The PA compares eigenvalues from the actual data (solid line) to those obtained from resampled data (dashed line) or simulated data (dotted line, not visible). Only one component could be extracted in this case because the PC resampled data line remained *above* the PC actual data line for the second component (as indicated by the circle ○).

The PCA was performed on selected WM measures, with attention to the first component, as suggested by the PA. This component accounted for 34.59% of the variance. Table 4 presents the component loadings λ, revealing that three of the eight WM measures (corsi-back_span, amma_total and corsi-for_span) fell below the typical loading threshold $|\lambda| > .30$ for the first component, indicating no meaningful contribution to the first component. However, these measures all demonstrated substantial loadings ($|\lambda| > .50$) on the second component. Indeed, once these WM measures were removed from the model, the overall loadings for the remaining variables slightly increased, resulting in an overall explained variance of 47.11% for the first component (also see Table 4).

Table 4

Results of the PCA

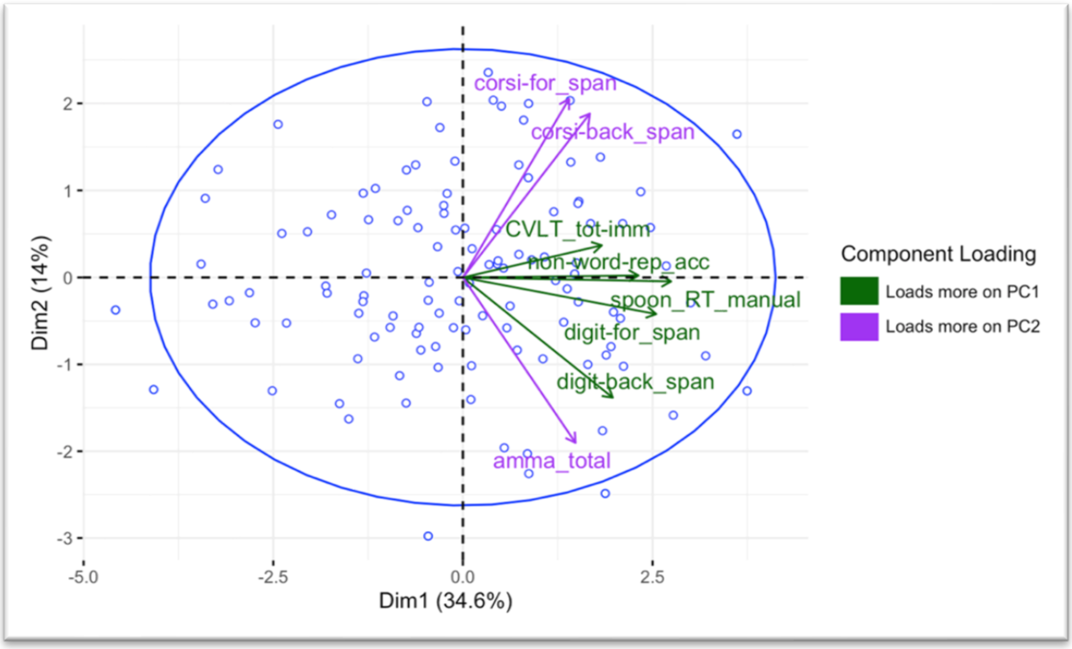
WM measure	PC1	PC2	PC 1
CVLT_tot-imm	0.32	0.10	0.32
digit-for_span	0.44	–0.11	0.48
digit-back_span	0.34	–0.37	0.42
corsi-back_span	0.29	0.51	—
corsi-for_span	0.24	0.56	—
non-word-rep_acc	0.40	0.01	0.48
spoon_RT_manual	0.47	–0.01	0.52
amma_total	0.26	–0.52	—
Cumulative explained variance	0.35	0.47	0.47

Note. WM measures with their respective loadings on the first two components (PC1 and PC2). After removal of WM measures with $|\lambda| \leq 0.30$, loadings onto the first component yielded a marginal increase.

Finally, it should be noted that the direction of some loadings varied depending on the component. This observation underpins that each component represents slightly different dimensional entities (Figure 6).

Figure 6

PCA Biplot Before Removal of Variables with $|\lambda| \leq 0.30$



With respect to the potential associations between the overall degree of multilingualism, the $FoLA_{ij}$-weighted, and $FoLA_{ij}$-filtered degrees of multilingualism, GAMs and LRMs were fitted. Since all LRMs yielded almost identical results as GAMs with no difference in significance, only LRMs are reported. To ensure that all predictors, including control variables, were not too highly correlated and to reduce the risk of inflated standard errors, which could render the regression coefficients unreliable (Field, 2012), multicollinearity was assessed using the variance inflation factor (VIF). The VIF measures the extent to which a predictor is correlated with other predictors in a LRM, helping to detect multicollinearity issues (Marcoulides & Raykov, 2019). All LRMs (apart from threshold nine-informal) were well below the commonly accepted threshold of $VIF \leq 5$ (O'Brien, 2007).

The overall degree of multilingualism (i.e., entropy for all languages irrespective of formality of acquisition) did not show any significant linear association with the WM component extracted, $F(4{,}110) = 0.28, p = .89$, with $R^2 = 0.001$ and adjusted $R^2 = -0.03$. The $FoLA_{ij}$-weighted degree of multilingualism based on informally acquired languages did not show a significant effect on the WM component, $F = (4{,}109) = 0.23, p = .92$, with $R^2 = 0.01$ and adjusted $R^2 = -0.03$. For formally acquired languages, the results were also non-significant, $F = (4{,}109) = 0.41, p = .80$, with $R^2 = 0.02$ and adjusted $R^2 = -0.02$.

These findings imply that the formality of language acquisition, when conceptualized continuously, is not linked to WM performance.

When the degree of multilingualism was calculated using manually filtered $FoLA_{ij}$ thresholds, the mid-range thresholds four and five yielded to be the clearest discrimination points for the relevance of formal vs. informal language acquisition to WM. In this case, the degree of multilingualism showed a significant association with WM performance for both formally ($p < .05$) and informally ($p < .01$) learnt languages. Notably, a threshold of three showed a marginally significant association ($p = .0483$) between the degree of multilingualism from informally learned languages and WM performance.

Since all data has been previously z-standardized, the original B effect sizes can be interpreted as equivalent to standardized β effect sizes, indicating the strength and direction of effects in standard deviation units. Across all thresholds, a clear pattern emerged: Formal language acquisition was consistently associated with positive effects on WM performance, while informal language acquisition was associated with negative effects on WM performance. For the significant thresholds, effect sizes ranged from $B/\beta = 0.38$ to $B/\beta = -1.01$, reflecting a medium to exceptionally large magnitude (Cohen, 1988). Comprehensive results are presented in Table 5 and Figure 7.

Table 5

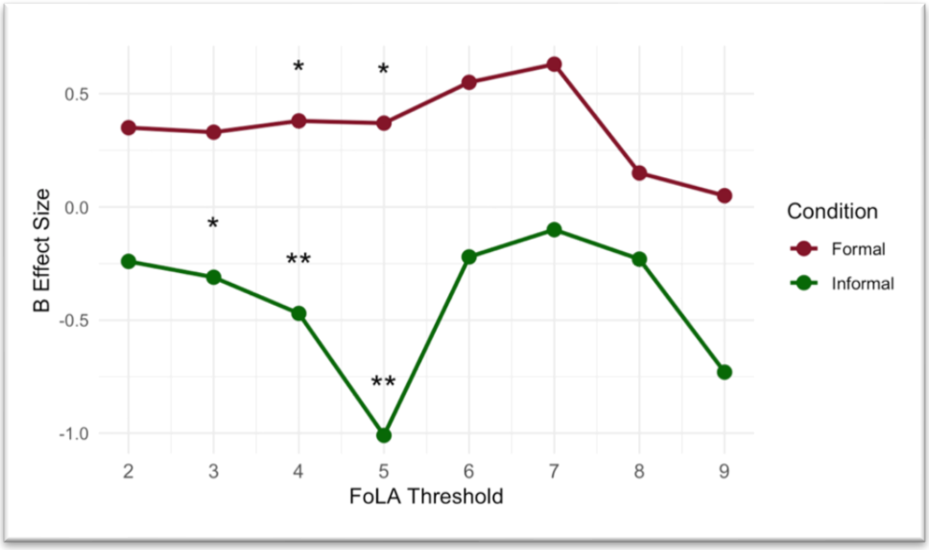
Overview of Relevant Terms in LRMs: Entropy Scores from Filtered $FoLA_{ij}$ Values

Formality	Threshold	*p*	*B*
FOR	2	0.1406	0.35
INF		0.1191	–0.24
FOR	3	0.0914	0.33
INF		0.0483*	–0.31
FOR	4	0.0251*	0.38
INF		0.0034**	–0.47
FOR	5	0.0125*	0.37
INF		0.0012**	–1.01
FOR	6	0.1009	0.55
INF		0.2080	–0.22
FOR	7	0.0788	0.63
INF		0.6398	–0.10
FOR	8	0.7330	0.15
INF		0.5908	–0.23
FOR	9	0.9251	0.05
INF		0.3887	–0.73

Note. FOR = formal language acquisition; INF = informal language acquisition.
*$p < .05$. **$p < .01$.

Figure 7

B Effect Sizes for Degree of Multilingualism (Formal vs. Informal Acquisition) on WM Performance



Note. *$p < .05$. **$p < .01$.

Discussion

In this study, it was hypothesized that the degree of multilingualism positively relates to (numerical-verbal) WM performance, with formal and informal language learning contributing differently to WM performance. Finally, findings will be discussed and placed into context.

H1: Different WM tasks correlate at least moderately ($r \geq 0.30$; Cohen, 1988) with each other.

For the first hypothesis (H1), 20 out of 28 correlations were significant (71.43%), with nine of these 20 significant coefficients (45%) reaching $r \geq 0.30$, suggesting that H1 can only be partially accepted. Although not all WM tasks correlated significantly with each other, they still largely represented a cohesive underlying construct, as indicated by the sufficient overall MSA index of .74, which could be captured through PCA. This does not necessarily undermine the validity of the findings, as they align with the idea that WM is not a monolithic construct but rather a set of interrelated yet distinct functions and processes (e.g., Repovš & Baddeley, 2006; Logie, 2011).

Of note, in psychology, Cohen's (1988) thresholds are often seen as too conservative, as most associations and effects are smaller yet still relevant. Applying updated thresholds, like $r \geq 0.20$ as suggested by Gignac and Szodorai (2016) or Funder and Ozer (2019), would reclassify 14 of the 20 significant correlations (70%) as medium-sized, offering a more context-appropriate interpretation.

H2: The different WM tasks can be clustered into at least two compound measures, i.e., components of WM: numerical-verbal and visuospatial.

Unlike Waris et al. (2017) a second component could not be extracted because the eigenvalue for the second component fell below the PC resampled threshold, as confirmed by the PA. This was most likely due to the limited sample size ($N = 116$), as many researchers recommend larger samples, especially when attempting to investigate latent psychological constructs like WM (Costello & Osborne, 2005; Kyriazos, 2018). With a larger sample, more robust methods like exploratory and confirmatory factor analyses would have been preferable (Watkins, 2018). To note, in Figure 5, the FA resampled line crosses the second component, implying that a factorial solution may have resulted in a two-factor WM structure. Despite this, the common two-component structure of WM was still apparent in the data, as the CBT and AMMA variables were excluded from the global WM component due to low component loadings. Nonetheless, these variables all loaded onto a second, unextracted component that seemed to reflect spatial ability. Applying the Kaiser criterion (Kaiser, 1960), i.e., extracting all components with eigenvalues over one, would have resulted in the extraction of two components. However, the more reliable PA and extraction based on resampled eigenvalues was chosen (Horn, 1965). Thus, in the present study, the global WM component essentially represented the *numerical-verbal* WM component.

H3: The degree of multilingualism is a stronger predictor of performance in numerical-verbal WM than in visuospatial WM.

No significant positive association was found between the degree of multilingualism and numerical-verbal WM, contrary to previous findings (Cockcroft et al., 2019; Espi-Sanchis & Cockcroft, 2022; Fyndanis et al., 2023). This suggests that the spread and level of language proficiencies across an individual's spoken languages are not directly linked to numerical-verbal WM performance, as underpinned by a negligible effect size of $B = 0.088$. This was unexpected, particularly since attention-heavy tasks like the backward digit span task, where multilinguals tend to excel (Achaa-Amankwaa et al., 2023; Cockcroft et al., 2019), were included in the numerical-verbal WM component analyzed. Despite the inclusion of language-heavy WM tasks, such as the Spoonerism task, NWRT, and CVLT, which theoretically should have favored individuals with higher proficiency across their spoken languages, this effect was not observed. This contrasts with the findings of Espi-Sanchis and Cockcroft (2022), who found improved performance in balanced multilinguals.

One explanation for this surprising finding could be the task- and age-dependency often observed in the bilingual advantage, which might also extend to multilingualism. For example, Ware et al. (2020) found that the effect of bilingualism on EFs is moderated by both age and task type, with consistent benefits only seen in specific tasks like the Attentional Network Task (ANT; de Souza Almeida et al., 2021). Since highly language-specific WM tasks were used that have not been previously examined in this context, it is possible that the multilingual advantage does not extend to these particular tasks. As Cockcroft et al. (2019, p. 18) pointed out, a "verbal disadvantage"—when vocabulary is not controlled—has been previously reported in bilinguals (see Table 2). The more languages one speaks, the less frequently each language is used, resulting in slower lexical access (Schmidtke, 2016; Shook et al., 2015; Weber & Cutler, 2004), and this effect may be even greater in multilinguals. Considering that the present sample included polyglots speaking up to 13 languages, with a median of five languages—far exceeding other studies where participants spoke up to six languages—such an effect may be more prominent in language-specific WM tasks, diluting the WM compound score. With regards to age, it is understood that the bilingual advantage is more noticeable in children and older adults, while it is muted but still present in young adulthood (Bialystok, 2017; Schroeder & Marian, 2017), which is why age was controlled for in this study.

Another aspect to be considered is the *linguistic distance* of languages, i.e., how similar or different two languages are in terms of their structure, grammar, or vocabulary. For instance, Kepinska et al. (2023) found that the smaller the linguistic distance, the easier it was for children to transfer knowledge and skills between languages, improving their ability to produce words in the target language. In line with this, Radman et al. (2021) demonstrated that individuals speaking closely related languages achieved better domain-general cognitive control. Therefore, linguistic distance may be a confounding factor in the relationship between the degree of multilingualism and numerical-verbal WM, as it was not controlled for.

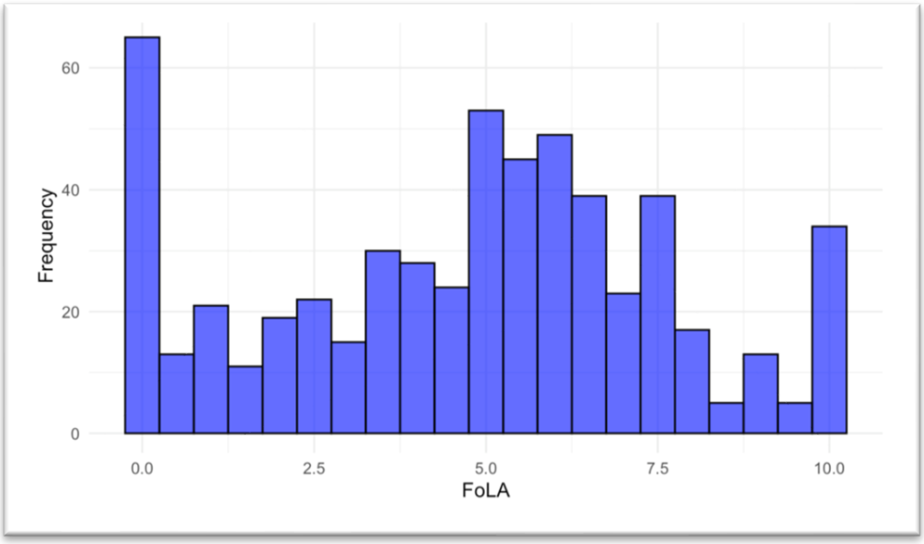
H4: The degree of multilingualism in formally acquired languages is associated with WM performance differently than in informally acquired languages.

No association was initially found between the degree of multilingualism and numerical-verbal WM when using weighted $FoLA_{ij}$ scores (*complementary approach*) for operationalizing formal and informal learning.

When setting manual $FoLA_{ij}$ thresholds (*segmented approach*; thresholds 2–9) to distinguish between formal and informal language acquisition, the mid-range cutoffs of four and five revealed significant, medium to exceptionally large effect sizes in opposite directions for numerical-verbal WM. These results imply that the distinction between formal and informal language acquisition was the most predictive for WM performance around this middle threshold range. Interestingly, the high concentration of $FoLA_{ij} = 0$ values, visible in Figure 8, did not shift the optimal threshold upwards as might be expected. A closer look at the distribution's low skewness (–0.13) and kurtosis (–0.82) implied that the visual clustering may have exaggerated the impact, justifying a mid-range threshold for $FoLA_{ij}$.

Figure 8

Histogram of the $FoLA_{ij}$ *Variable*



Moreover, the clustering near the center of the distribution suggested that many participants became multilingual through a blend of formal and informal learning. Notably, *formal* language acquisition produced *positive* effect sizes across all thresholds, while *informal* learning consistently yielded *negative* effects (see Table 5). The observed pattern suggests that structured formal learning, combined with higher and more balanced language proficiency across all spoken languages, benefited numerical-verbal WM, whereas the cognitive load of managing informally learnt languages hindered numerical-verbal WM performance.

The extreme effect of the degree of multilingualism in informal languages on WM performance at thresholds above five (to $B/\beta = -1.01$) warrants further investigation. From a descriptive perspective, 43.51% of $FoLA_{ij}$ values were below the threshold of five, and 56.49% were $FoLA_{ij} \geq 5$. As such, the threshold of five appeared to be especially impactful as it nearly bisected the data. This balanced division implied that values around this midpoint capture a pivotal range within the data distribution, potentially amplifying the contrast between formal vs. informal learning.

A potential explanation for the surprising negative effects of informal language acquisition on numerical-verbal WM could lie in *metalinguistic awareness* (Roehr-Brackin, 2018). In informal environments, one could expect that learners are more likely to immerse themselves unconsciously over time, without the focused engagement typical of formal settings like classrooms with teachers (who might help to maintain focus). Since key WM processes, such as subvocal rehearsal and attentional refreshing, are understood to be conscious, explicit strategies (Baddeley, 2003), informal language learning may lead to reduced WM involvement and less “attentional training”, resulting in overall lower WM performance. Christianson and Deshaies (2020) highlight that multilinguals who usually show high metalinguistic awareness often benefit more from explicit learning strategies, especially in unstructured environments, while implicit learning is typically linked to L1 acquisition. However, this argument would explain a null effect only and remains speculative, as unconscious processes are difficult to measure, and individual motivation plays a significant role, too (e.g., see L2 Motivational Self System Theory; Dörnei, 2005, 2009). For instance, a highly motivated language learner studying abroad may engage more consciously even in informal settings, influencing cognitive performance—something that is rarely controlled for in studies (Bialystok et al., 2012).

Another hypothetical reason for this negative effect of multilingualism in informally acquired languages could be that informal learning often lacks structured environments and reinforcement, potentially increasing the likelihood of so called *lexical interference* (Kroll et al., 2006). In these unstructured settings, learners may more likely unintentionally activate representations of words across multiple non-target languages, creating confusion and reducing language processing efficiency. Additionally, languages acquired informally may rely on less solidified mental representations, making them potentially more vulnerable to lexical interference. Given the uniquely multilingual sample in this study, juggling so many languages simultaneously might overwhelm the cognitive system (Schroeder & Marian, 2017), especially within the unstructured contexts typical of informal learning.

Furthermore, it is important to note that entropy was computed based on speaking competence rather than other language modalities like listening, reading, or writing. Fyndanis et al. (2023) argue that speaking and listening in real-time impose greater demands on WM compared to reading or writing because there is no opportunity to revisit the content and mitigate any WM limitations. Their study identified the number of foreign languages as the sole predictor of WMC. However, in the present case, learning (multiple) languages in a formal setting—typically more reliant on writing—might alleviate the burden on WM due to the ability to self-correct. This process likely leads to more gradual yet effective improvements in memory and language skills. In contrast, informal settings potentially impose a heavier burden on WM due to the need for spontaneous production. This may enhance fluency but might not contribute as much to long-term memory development, which could explain the direction of the effects found.

The positive effect of formal language learning on numerical-verbal WM performance could be understood through the Global Workspace Theory (GWT; Baars, 1988), which suggests that conscious processes are essential for organizing unconscious functions. As tasks become more practiced and automatic, they transition to unconscious specialized functions, potentially reducing the load on WM (Baars & Franklin, 2003). This shift could be often seen in classroom or mobile app settings, where learners engage in highly conscious, rule-based instructions and eventually internalize these. Similarly, Sweller's Cognitive Load Theory (Sweller, 1994) argues that structured, instructional learning environments help manage complex tasks by breaking them into smaller, manageable chunks, thus preventing cognitive overload and maximizing WM efficiency, which would explain the positive trend of formal language acquisition on numerical-verbal WM. As a remark, it is worth noting that while the conscious nature of WM is well-established, some preliminary findings indicate potential unconscious features in WM that call for additional research (Trübutschek et al., 2017).

Conclusion

This study investigated how multilingualism impacts WM performance in a highly multilingual sample. Unlike prior research, no association between the overall degree of multilingualism and numerical-verbal WM was found. A newly introduced variable, $FoLA_{ij}$, proved to be effective in distinguishing between the effect of formally and informally learned languages on WM, with key thresholds of four and five on its scale. Defining language (in-)formality in a continuous manner did not yield significant results. The most surprising finding suggested that informal language acquisition may be negatively related to numerical-verbal WM performance, whereas formal learning may have a positive relation.

Methodological limitations, including the limited sample size and unaccounted factors like linguistic distance, vocabulary, motivation, and socio-economic status, prevent broad generalization at this point in time. Future studies should address these factors, consider additional language modalities, and develop more elaborate tools to measure the formality of language acquisition. Moreover, investigating the electiveness of language learning could shed light on more contextual factors influencing cognition. Additionally, combining these findings with structural brain analyses could clarify whether behavioral patterns mirror anatomical changes. Although none of the hypotheses posited could be fully accepted, reporting these results is crucial for addressing the publication bias often seen in bilingualism research. The field of multilingualism and its link to EFs remains a valuable avenue for future research.

References

Abutalebi, J., Annoni, J. M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., Lazeyras, F., Cappa, S. F., & Khateb, A. (2008). Language control and lexical competition in bilinguals: An event-related fMRI study. *Cerebral Cortex*, *18*(7), 1496–1505. https://doi.org/10.1093/cercor/bhm182

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20*(3), 242–275. https://doi.org/https://doi.org/10.1016/j.jneuroling.2006.10.003

Achaa-Amankwaa, P., Kushnereva, E., Miksch, H., Stumme, J., Heim, S., & Ebersbach, M. (2023). Multilingualism is associated with small task-specific advantages in cognitive performance of older adults. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-43961-7

Adams, E. J., Nguyen, A. T., & Cowan, N. (2018). Theories of working memory: Differences in definition, degree of modularity, role of attention, and purpose. *Language, Speech, and Hearing Services in Schools*, *49*(3), 340–355. https://doi.org/10.1044/2018_lshss-17-0114

Adesope, O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research*, *80*, 207–245. https://doi.org/10.3102/0034654310368803

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, *14*(5), 255–259. https://doi.org/10.1111/j.0963-7214.2005.00376.x

Antón, E., Carreiras, M., & Duñabeitia, J. A. (2019). The impact of bilingualism on executive functions and working memory in young adults. *PLOS ONE*, *14*(2), e0206770. https://doi.org/10.1371/journal.pone.0206770

Antoniou, M. (2019). The advantages of bilingualism debate. *Annual Review of Linguistics*, *5*(1), 395–415. https://doi.org/10.1146/annurev-linguistics-011718-011820

Arce, T., & McMullen, K. (2021). The Corsi block-tapping test: Evaluating methodological practices with an eye towards modern digital frameworks. *Computers in Human Behavior Reports*, *4*. https://doi.org/10.1016/j.chbr.2021.100099

Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge University Press.

Baars, B. J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Sciences*, *7*(4), 166–172. https://doi.org/10.1016/s1364-6613(03)00056-1

Baayen, R. H., & Linke, M. (2020). Generalized additive mixed models. In M. Paquot & S. T. Gries (Eds.), *A practical handbook of corpus linguistics* (pp. 563–591). Springer International Publishing. https://doi.org/10.1007/978-3-030-46216-1_23

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A. (2010). Working memory. *Current Biology*, *20*, 136–140. https://doi.org/10.1016/j.cub.2009.12.014

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. (2015). Working Memory in Second Language Learning. In W. Zhisheng, M. Mailce Borges, & M. Arthur (Eds.), *Working memory in second language acquisition and processing* (pp. 17–28). Multilingual Matters. https://doi.org/doi:10.21832/9781783093595-005

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/https://doi.org/10.1016/S0079-7421(08)60452-1

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, *130*(4), 553–573. https://doi.org/10.1037/0033-2909.130.4.553

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, *133*(1), 83–100. https://doi.org/10.1037/0096-3445.133.1.83

Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. Psychology Press.

Bates, D. M. M. B. B., & Walker, S. (2015). *lme4: Linear Mixed-Effects Models Using 'Eigen' and S4*. R package version 1.1-35.3. https://CRAN.R-project.org/package=lme4

Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives*, *9*(2), 117–121. https://doi.org/10.1111/cdep.12116

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, *143*(3), 233–262. https://doi.org/10.1037/bul0000099

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*(2), 290–303. https://doi.org/10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences*, *16*(4), 240–250. https://doi.org/10.1016/j.tics.2012.03.001

Bialystok, E., & Craik, F. I. M. (2022). How does bilingualism modify cognitive function? Attention to the mechanism. *Psychonomic Bulletin and Review*, *29*(4), 1246–1269. https://doi.org/10.3758/s13423-022-02057-5

Bialystok, E., Poarch, G., Luo, L., & Craik, F. I. M. (2014). Effects of bilingualism and aging on executive function and working memory. *Psychology and Aging*, *29*(3), 696–705. https://doi.org/10.1037/a0037254

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish–Dutch children. *Journal of Experimental Child Psychology*, *128*, 105–119. https://doi.org/https://doi.org/10.1016/j.jecp.2014.06.007

Bonifacci, P., Giombini, L., Bellocchi, S., & Contento, S. (2011). Speed of processing, anticipation, inhibition and working memory in bilinguals. *Developmental Science*, *14*(2), 256–269. https://doi.org/10.1111/j.1467-7687.2010.00974.x

Brito, N., Sebastian Galles, N., & Barr, R. (2014). Differences in language exposure and its effects on memory flexibility in monolingual, bilingual, and trilingual infants. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728914000789

Brunetti, R., Del Gatto, C., & Delogu, F. (2014). eCorsi: Implementation and testing of the Corsi block-tapping task for digital tablets. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00939

Buchsbaum, B. R., & D'Esposito, M. (2008). The search for the phonological store: From loop to convolution. *Journal of Cognitive Neuroscience*, *20*(5), 762–778. https://doi.org/10.1162/jocn.2008.20501

Calvo, N., Ibáñez, A., & García, A. M. (2016). The Impact of Bilingualism on Working Memory: A Null Effect on the Whole May Not Be So on the Parts. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00265

Chertkow, H., Whitehead, V., Phillips, N., Wolfson, C., Atherton, J., & Bergman, H. (2010). Multilingualism (but not always bilingualism) delays the onset of Alzheimer disease:

Evidence from a bilingual community. *Alzheimer Disease and Associated Disorders*, *24*(2), 118–125. https://doi.org/10.1097/WAD.0b013e3181ca1221

Christianson, K., & Deshaies, S.-E. (2020). Learning languages in informal environments: Some cognitive considerations. In *The handbook of informal language learning* (pp. 27–37). Wiley Blackwell. https://doi.org/10.1002/9781119472384.ch2

Cockcroft, K., Wigdorowitz, M., & Liversage, L. (2019). A multilingual advantage in the components of working memory. *Bilingualism: Language and Cognition*, *22*(1), 15–29. https://doi.org/10.1017/s1366728917000475

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. In.

Comishen, K. J., & Bialystok, E. (2021). Increases in attentional demands are associated with language group differences in working memory performance. *Brain and Cognition*, *147*. https://doi.org/10.1016/j.bandc.2020.105658

Comishen, K. J., Bialystok, E., & Adler, S. A. (2019). The impact of bilingual environments on selective attention in infancy. *Developmental Science*, *22*(4), e12797. https://doi.org/10.1111/desc.12797

Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain* McGill University].

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*.

Cowan, N. (1998). Visual and auditory working memory capacity. *Trends in Cognitive Sciences*, *2*, 77. https://doi.org/10.1016/S1364-6613(98)01144-9

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin and Review*, *24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Crawford, A. V., Green, S. B., Levy, R., Lo, W.-J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement*, *70*(6), 885–901. https://doi.org/10.1177/0013164410379332

D'Souza, A. A., Moradzadeh, L., & Wiseheart, M. (2018). Musical training, bilingualism, and executive function: Working memory and inhibitory control. *Cognitive Research: Principles and Implications*, *3*(1). https://doi.org/10.1186/s41235-018-0095-6

Davidson, C., Shing, Y. L., McKay, C., Rafetseder, E., & Wijeakumar, S. (2023). The first year in formal schooling improves working memory and academic abilities.

Developmental Cognitive Neuroscience, *60*, 101205. https://doi.org/https://doi.org/10.1016/j.dcn.2023.101205

de Bruin, A., Treccani, B., & Sala, S. D. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science*, *26*, 107. https://doi.org/10.1177/0956797614557866

de Souza Almeida, R., Faria-Jr, A., & Klein, R. M. (2021). On the origins and evolution of the Attention Network Tests. *Neuroscience & Biobehavioral Reviews*, *126*, 560–572. https://doi.org/https://doi.org/10.1016/j.neubiorev.2021.02.028

De Wilde, V., Brysbaert, M., & Eyckmans, J. (2022). Formal versus informal L2 learning: How do individual differences and word-related variables influence French and English l2 vocabulary learning in Dutch-speaking children? *Studies in Second Language Acquisition*, *44*(1), 87–111. https://doi.org/10.1017/S0272263121000097

Delis, D., Kramer, J., Kaplan, E., & Ober, B. (1987). *California verbal learning test research edition manual*. San Antonio: The Psychological Corporation.

Delis, D. C., Cullum, C. M., Butters, N., Cairns, P., & Prifitera, A. (1988). Wechsler memory scale-revised and california verbal learning test: Convergence and divergence. *Clinical Neuropsychologist*, *2*(2), 188–196. https://doi.org/10.1080/13854048808520100

Deweer, B., Poitrenaud, J., Kalafat, M., & Van der Linden, M. (2008). *CVLT – Test d'apprentissage et de mémoire verbale*. ECPA – Pearson Assessments. https://www.pearsonclinical.fr/cvlt-test-dapprentissage-et-de-memoire-verbale

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Diamond, J. (2010). The benefits of multilingualism. *Science*, *330*(6002), 332–333. https://doi.org/10.1126/science.1195067

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. https://doi.org/10.3758/s13423-019-01567-z

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Taylor & Francis Group. https://ebookcentral.proquest.com/lib/univie/detail.action?docID=257290

Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Multilingual Matters.

Dressman, M., & Sadler, R. (2020). *The handbook of informal language learning*. Wiley Blackwell. https://doi.org/10.1002/9781119472384

https://ubdata.univie.ac.at/AC15618019

Engel De Abreu, P. M. J. (2011). Working memory in multilingual children: Is there a bilingual effect? *Memory*, *19*(5), 529–537. https://doi.org/10.1080/09658211.2011.590504

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23. https://doi.org/10.1111/1467-8721.00160

Espi-Sanchis, G., & Cockcroft, K. (2022). Working memory and multilingualism: Balanced language proficiency predicts verbal working memory. *International Journal of Bilingual Education and Bilingualism*, *25*, 2976–2990. https://doi.org/10.1080/13670050.2021.1997901

Eurostat. (2019, October, 2024). *Foreign language skills statistics*. Retrieved September 9, 2024 from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Foreign_language_skills_statistics#Number_of_foreign_languages_known

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Field, A. P. (2012). *Discovering statistics using R*. SAGE Publications. https://app.talis.com/textbooks/9781446258460

Finch, J. E. (2019). Do schools promote executive functions? Differential working memory growth across school-year and summer months. *AERA Open*, *5*(2), 2332858419848443. https://doi.org/10.1177/2332858419848443

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Fyndanis, V., Cameron, S., Hansen, P. B., Norvik, M. I., & Simonsen, H. G. (2023). Multilingualism and verbal short-term/working memory: Evidence from academics. *Bilingualism: Language and Cognition*, *26*(3), 490–503. https://doi.org/10.1017/s1366728922000621

Gajewski, P. D., Hanisch, E., Falkenstein, M., Thönes, S., & Wascher, E. (2018). What does the n-Back task measure as we get older? Relations between working-memory measures and other cognitive functions across the lifespan. *Frontiers in Psychology*, *9*, Article 2208. https://doi.org/10.3389/fpsyg.2018.02208

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*(2), 103–127. https://doi.org/10.1080/09658219408258940

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and individual differences*, *102*, 78. https://doi.org/10.1016/j.paid.2016.06.069

Gordon, E. E. (1989). *Music Aptitude and Related Tests. An Introduction*. GIA Music. www.giamusic.com

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

Grundy, J., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second language Research*, *33*, 325–340. https://doi.org/10.1177/0267658316678286

Grundy, J. G. (2020). The effects of bilingualism on executive functions: An updated quantitative analysis. *Journal of Cultural Cognitive Science*, *4*(2), 177–199. https://doi.org/10.1007/s41809-020-00062-5

Gullifer, J. W., Kousaie, S., Gilbert, A. C., Grant, A., Giroud, N., Coulter, K., Klein, D., Baum, S., Phillips, N., & Titone, D. (2021). Bilingual language experience as a multidimensional spectrum: Associations with objective and subjective language proficiency. *Applied Psycholinguistics*, *42*(2), 245–278. https://doi.org/10.1017/s0142716420000521

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition*, *23*(2), 283–294. https://doi.org/10.1017/s1366728919000026

Gullifer, J. W., & Titone, D. (2021a). Engaging proactive control: Influences of diverse language experiences using insights from machine learning. *Journal of Experimental Psychology: General*, *150*(3), 414–430. https://doi.org/10.1037/xge0000933

Gullifer, J. W., & Titone, D. (2021b). Engaging proactive control: Influences of diverse language experiences using insights from machine learning. *Journal of Experimental Psychology*, *150*(3), 414–430. https://doi.org/10.1037/xge0000933

Hager, P. J. (2012). Formal learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 1314–1316). Springer US. https://doi.org/10.1007/978-1-4419-1428-6_160

Harley, T. A. (2006). Speech errors: Psycholinguistic approach. In K. Brown (Ed.), *Encyclopedia of language & linguistics (Second Edition)* (pp. 739–745). Elsevier. https://doi.org/https://doi.org/10.1016/B0-08-044854-2/00797-5

Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Science*, *1*(3), 297–310. https://doi.org/10.1214/ss/1177013604

Hausser, J., & Strimmer, K. (2021). *entropy: Estimation of entropy*. R package version 1.3.1. https://CRAN.R-project.org/package=entropy

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179–185. https://doi.org/10.1007/bf02289447

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, *24*, 417–441. https://doi.org/10.1037/h0071325

Inayati, N., Sanjani, M. I., Jayanti, F. G., Gao, X., & Nguyen, H. T. M. (2024). Online informal language learning (OILL): A systematic review of studies (2014–2023). *Innovation in Language Learning and Teaching*, 1–17. https://doi.org/10.1080/17501229.2024.2355296

Janus, M., & Bialystok, E. (2018). Working memory with emotional distraction in monolingual and bilingual children. *Frontiers in Psychology*, *9*, Article 1582. https://doi.org/10.3389/fpsyg.2018.01582

Johnson, M., & Majewska, D. (2024). *Formal, non-formal, and informal learning: What are they, and how can we research them?* C. U. P. Assessment. https://www.cambridgeassessment.org.uk/Images/665425-formal-non-formal-and-informal-learning-what-are-they-and-how-can-we-research-them-.pdf

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 615–622. https://doi.org/10.1037/0278-7393.33.3.615

Kavé, G., Eyal, N., Shorek, A., & Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychology and Aging*, *23*(1), 70–78. https://doi.org/10.1037/0882-7974.23.1.70

Kepinska, O., Caballero, J., Oliver, M., Marks, R. A., Haft, S. L., Zekelman, L., Kovelman, I., Uchikoshi, Y., & Hoeft, F. (2023). Language combinations of multilinguals are reflected in their first-language knowledge and processing. *Scientific Reports*, *13*(1), Article 1947. https://doi.org/10.1038/s41598-023-27952-2

Kepinska, O., Dalboni Da Rocha, J., Tuerk, C., Hervais-Adelman, A., Bouhali, F., Green, D., Price, C. J., & Golestani, N. (2023). Auditory cortex anatomy reflects multilingual phonological experience. *eLife*. https://doi.org/10.7554/eLife.90269.1

Kessels, R. P. C., van Zandvoort, M. J. E., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000). The Corsi block-tapping task: Standardization and normative sata. *Applied Neuropsychology*, *7*(4), 252–258. https://doi.org/10.1207/S15324826AN0704_8

Knoop-van Campen, C. A. N., Segers, E., & Verhoeven, L. (2018). How phonological awareness mediates the relation between working memory and word reading efficiency in children with dyslexia. *Dyslexia*, *24*(2), 156–169. https://doi.org/https://doi.org/10.1002/dys.1583

Korkmaz, S. G. D., & Zararsiz, G. (2014). *MVN: Multivariate Normality Tests*. R package version 5.9. https://CRAN.R-project.org/package=MVN

Kosmidis, M. H., Zafiri, M., & Politimou, N. (2011). Literacy versus formal schooling: Influence on working memory. *Archives of Clinical Neuropsychology*, *26*(7), 575–582. https://doi.org/10.1093/arclin/acr063

Kotyusov, A. I., Kasanov, D., Kosachenko, A. I., Gashkova, A. S., Pavlov, Y. G., & Malykh, S. (2023). Working memory capacity depends on attention control, but not selective attention. *Behavioral Sciences*, *13*(2), 92. https://doi.org/10.3390/bs13020092

Krashen, S. D. (1976). Formal and informal linguistic environments in language acquisition and language learning. *TESOL Quarterly*, *10*(2), 157–168. https://doi.org/10.2307/3585637

Kroll, J., Bobb, S., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech.

Bilingualism: Language and Cognition, *9*(2), 119–135. https://doi.org/10.1017/S1366728906002483

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, *128*(3), 416–430. https://doi.org/10.1016/j.actpsy.2008.02.001

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics*, *1*, 377–394. https://doi.org/10.1146/annurev-linguist-030514-124937

Kroll, J. F., Dussias, P. E., Bogulski, C. A., & Kroff, J. R. V. (2012). Juggling two languages in one mind: What bilinguals tell us about language processing and its consequences for cognition. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 229–262). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-394393-4.00007-8

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*(8), 2207–2230. https://doi.org/10.4236/psych.2018.98126

Lacritz, L. H., & Cullum, C. M. (1998). The Hopkins Verbal Learning Test and CVLT: A preliminary comparison. *Archives of Clinical Neuropsychology*, *13*(7), 623–628.

Ladowsky-Brooks, R. L. (2016). Four-hour delayed memory recall for stories: Theoretical and clinical implications of measuring accelerated long-term forgetting. *Applied Neuropsychology: Adult*, *23*(3), 205–212. https://doi.org/10.1080/23279095.2015.1030670

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. https://doi.org/10.1037/bul0000142

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Liu, Q., & Liu, Z. (2021). A review on the relationship between bilingualism and working memory. *Open Journal of Modern Linguistics*, *11*, 121–134. https://doi.org/10.4236/ojml.2021.112010

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science*, *20*(4), 240–245. https://doi.org/10.1177/0963721411415340

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*(5), 605–621. https://doi.org/10.1080/20445911.2013.795574

Lukasik, K. M., Lehtonen, M., Soveri, A., Waris, O., Jylkkä, J., & Laine, M. (2018). Bilingualism and working memory performance: Evidence from a large-scale online study. *PLOS ONE*, *13*(11), e0205916. https://doi.org/10.1371/journal.pone.0205916

Luo, L., Craik, F. I. M., Moreno, S., & Bialystok, E. (2013). Bilingualism interacts with domain in a working memory task: Evidence from aging. *Psychology and aging*, *28*, 28–34. https://doi.org/10.1037/a0030875

Majerus, S., & Van der Linden, M. (2003). Long-term memory effects on verbal short-term memory: A replication study. *British Journal of Developmental Psychology*, *21*(2), 303–310. https://doi.org/10.1348/026151003765264101

Marcoulides, K. M., & Raykov, T. (2019). Evaluation of variance inflation factors in regression models using latent variable modeling methods. *Educational and Psychological Measurement*, *79*(5), 874–882. https://doi.org/10.1177/0013164418817803

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*, 519–530. https://doi.org/10.1093/biomet/57.3.519

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *J Speech Lang Hear Res*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222–243. https://doi.org/10.1037/a0017619

McGivney, V. (2001). Informal learning: A neglected species. *Journal of Adult and Continuing Education*, *7*(2), 101–107. https://doi.org/10.1177/147797140100700207

Milner, B. (1971). Interhemispheric differences in the localization of psychological processes in man. *British Medical Bulletin*, *27*(3), 272–277. https://doi.org/10.1093/oxfordjournals.bmb.a070866

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. https://doi.org/10.1006/cogp.1999.0734

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*(2), 187–202. https://doi.org/10.1016/j.jecp.2012.09.002

Morrison, C., Kamal, F., & Taler, V. (2019). The influence of bilingualism on working memory event-related potentials. *Bilingualism*, *22*, 191–199. https://doi.org/10.1017/S1366728918000391

Namazi, M., & Thordardottir, E. (2010). A working memory, not bilingual advantage, in controlled attention. *International Journal of Bilingual Education and Bilingualism*, *13*, 597–616. https://doi.org/10.1080/13670050.2010.488288

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*, 673–690. https://doi.org/10.1007/s11135-006-9018-6

Paap, K., Myuz, H., Anders-Jefferson, R., Bockelman, M., Mikulinsky, R., & Sawi, O. (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. https://doi.org/10.1080/20445911.2016.1248436

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. https://doi.org/10.1016/j.cogpsych.2012.12.002

Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, *69*, 265–278. https://doi.org/10.1016/j.cortex.2015.04.014

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh and Dublin philosophical magazine and journal of science*, *2*(11), 559–572. https://doi.org/10.1080/14786440109462720

Perquin, M., Vaillant, M., Schuller, A.-M., Pastore, J., Dartigues, J.-F., Lair, M.-L., & Diederich, N. (2013). Lifelong exposure to multilingualism: New evidence to support cognitive reserve hypothesis. *PLoS ONE*, *8*(4), e62030. https://doi.org/10.1371/journal.pone.0062030

Platz, F., Kopiez, R., Lehmann, A. C., & Wolf, A. (2022). Measuring audiation or tonal memory? Evaluation of the discriminant validity of Edwin E. Gordon's "Advanced Measures of Music Audiation". *Music & Science*, *5*, 1–16. https://doi.org/10.1177/20592043221105270

Poarch, G. J., & van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *113*, 535–551. https://doi.org/10.1016/j.jecp.2012.06.013

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. Version 4.1.1. https://www.R-project.org/

Radman, N., Jost, L., Dorood, S., Mancini, C., & Annoni, J.-M. (2021). Language distance modulates cognitive control in bilinguals. *Scientific Reports*, *11*, Article 24131. https://doi.org/10.1038/s41598-021-02973-x

Rampinini, A., Balboni, I., Golestani, N., & Berthele, R. (2024). A behavioural exploration of language aptitude and experience, cognition and more using Graph Analysis. *Brain Research*, *1842*, Article 149109. https://doi.org/https://doi.org/10.1016/j.brainres.2024.149109

Ramus, F. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain*, *126*(4), 841–865. https://doi.org/10.1093/brain/awg076

Ratiu, I., & Azuma, T. (2015). Working memory capacity: Is there a bilingual advantage? *Journal of Cognitive Psychology*, *27*(1), 1–11. https://doi.org/10.1080/20445911.2014.976226

Repovš, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience*, *139*, 5–21. https://doi.org/10.1016/j.neuroscience.2005.12.061

Revelle, W. (2024). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University, Evanston, Illinois. https://CRAN.R-project.org/package=psych

Richardson, F. M., Ramsden, S., Ellis, C., Burnett, S., Megnin, O., Catmur, C., Schofield, T. M., Leff, A. P., & Price, C. J. (2011). Auditory short-term memory capacity correlates with gray matter density in the left posterior STS in cognitively normal and dyslexic adults. *Journal of Cognitive Neuroscience*, *23*(12), 3746–3756. https://doi.org/10.1162/jocn_a_00060

Richardson, J. T. E. (2007). Measures of short-term memory: A historical review. *Cortex*, *43*(5), 635–650. https://doi.org/https://doi.org/10.1016/S0010-9452(08)70493-3

Riffenburgh, R. H. (2012). Chapter 28 - Methods you might meet, but not every day. In R. H. Riffenburgh (Ed.), *Statistics in Medicine* (3 ed., pp. 581–591). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-384864-2.00028-7

Roberts, G., Quach, J., Mensah, F., Gathercole, S., Gold, L., Anderson, P., Spencer-Smith, M., & Wake, M. (2015). Schooling Duration Rather Than Chronological Age Predicts Working Memory Between 6 and 7 Years: Memory Maestros Study. *Journal of Developmental & Behavioral Pediatrics*, *36*(2). https://journals.lww.com/jrnldbp/fulltext/2015/02000/schooling_duration_rather_than_chronological_age.4.aspx

Roehr-Brackin, K. (2018). Introduction. In *Metalinguistic awareness and second language acquisition* (pp. 1–4). Routledge. https://doi.org/10.4324/9781315661001

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*, Article 678. https://doi.org/10.3389/fpsyg.2016.00678

Schroeder, S. R., & Marian, V. (2017). Cognitive consequences of trilingualism. *International Journal of Bilingualism*, *21*(6), 754–773. https://doi.org/10.1177/1367006916637288

Seashore, C. E. (1919). *The psychology of musical talent* [doi:10.1037/13031-000]. Silver, Burdett & Company. https://doi.org/10.1037/13031-000

Shannon, C. E. (2001). A mathematical theory of communication. *Mobile Computing and Communications Review*, *5*(1), 3–55. https://doi.org/10.1145/584091.584093

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

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. https://doi.org/10.1007/s10936-014-9322-6

Siddi, S., Preti, A., Lara, E., Brébion, G., Vila, R., Iglesias, M., Cuevas-Esteban, J., López-Carrilero, R., Butjosa, A., & Haro, J. M. (2020). Comparison of the touch-screen and traditional versions of the Corsi block-tapping test in patients with psychosis and healthy controls. *BMC Psychiatry*, *20*(1), Article 329. https://doi.org/10.1186/s12888-020-02716-8

Souza-Talarico, J. N. D., Caramelli, P., Nitrini, R., & Chaves, E. C. (2007). The influence of schooling on working memory performance in elderly individuals without cognitive decline. *Dementia & Neuropsychologia*, *1*(3), 276–281. https://doi.org/10.1590/s1980-57642008dn10300009

Steger, M. F. (2006). An illustration of issues in factor extraction and identification of dimensionality in psychological assessment data. *Journal of Personality Assessment*, *86*(3), 263–272. https://doi.org/10.1207/s15327752jpa8603_03

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, *69*(347), 730–737. https://doi.org/10.1080/01621459.1974.10480196

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Szenkovits, G., & Ramus, F. (2005). Exploring dyslexics' phonological deficit I: Lexical vs sub-lexical and input vs output processes. *Dyslexia*, *11*(4), 253–268. https://doi.org/10.1002/dys.308

Tao, L., Wang, G., Zhu, M., & Cai, Q. (2021). Bilingualism and domain-general cognitive functions from a neural perspective: A systematic review. *Neuroscience & Biobehavioral Reviews*, *125*, 264–295. https://doi.org/https://doi.org/10.1016/j.neubiorev.2021.02.029

Tiv, M., Gullifer, J. W., Feng, R. Y., & Titone, D. (2020). Using network science to map what Montréal bilinguals talk about across languages and communicative contexts. *Journal of Neurolinguistics*, *56*, Article 100913. https://doi.org/10.1016/j.jneuroling.2020.100913

Trübutschek, D., Marti, S., Ojeda, A., King, J.-R., Mi, Y., Tsodyks, M., & Dehaene, S. (2017). A theory of working memory without consciousness or sustained activity. *eLife*, *6*, e23871. https://doi.org/10.7554/eLife.23871

Valls-Serrano, C., De Francisco, C., Vélez-Coto, M., & Caracuel, A. (2022). Visuospatial working memory and attention control make the difference between experts, regulars and non-players of the videogame League of Legends. *Frontiers in Human Neuroscience*, *16*, Article 933331. https://doi.org/10.3389/fnhum.2022.933331 van den Noort, M., Struys, E., Bosch, P., Jaswetz, L., Perriard, B., Yeo, S., Barisch, P., Vermeire, K., Lee, S.-H., & Lim, S. (2019). Does the bilingual advantage in cognitive control exist and if so, what are its modulating factors? A systematic review. *Behavioral Sciences*, *9*(3), 27. https://doi.org/10.3390/bs9030027

Varvara, P., Varuzza, C., Sorrentino, A. C. P., Vicari, S., & Menghini, D. (2014). Executive functions in developmental dyslexia. *Frontiers in Human Neuroscience*, *8*, Article 120. https://doi.org/10.3389/fnhum.2014.00120

Vaughan, R. S., & Laborde, S. (2021). Attention, working-memory control, working-memory capacity, and sport performance: The moderating role of athletic expertise. *European Journal of Sport Science*, *21*(2), 240–249. https://doi.org/10.1080/17461391.2020.1739143

Ware, A. T., Kirkovski, M., & Lum, J. A. G. (2020). Meta-analysis reveals a bilingual advantage that is dependent on task and age. *Frontiers in Psychology*, *11*, Article 1458. https://doi.org/10.3389/fpsyg.2020.01458

Waris, O., Soveri, A., Ahti, M., Hoffing, R. C., Ventus, D., Jaeggi, S. M., Seitz, A. R., & Laine, M. (2017). A latent factor analysis of working memory measures using large-scale data. *Frontiers in Psychology*, *8*, Articel 1062. https://doi.org/10.3389/fpsyg.2017.01062

Warmington, M. A., Kandru-Pothineni, S., & Hitch, G. J. (2019). Novel-word learning, executive control and working memory: A bilingual advantage. *Bilingualism: Language and Cognition*, *22*(4), 763–782. https://doi.org/10.1017/s136672891800041x

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, *44*(3), 219–246. https://doi.org/10.1177/0095798418771807

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25. https://doi.org/10.1016/S0749-596X(03)00105-0

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—Fourth Edition*. https://doi.org/https://doi.org/10.1037/t15169-000

Wei, T., & Simko, V. (2021). *corrplot: Visualization of a Correlation Matrix*. R package version 0.92. https://CRAN.R-project.org/package=corrplot

Wickham, H., & Henry, L. (2023). *purrr: Functional Programming Tools*. R package version 1.0.2. https://CRAN.R-project.org/package=purrr

Wickham, H. F. R. H. L., & Müller, K. (2023). *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4. https://CRAN.R-project.org/package=dplyr

Wickham, H. H. J., & Bryan, J. (2024). *readr: Read Rectangular Text Data*. R package version 2.1.5. https://CRAN.R-project.org/package=readr

Wickham, H. V. D., & Girlich, D. (2024). *tidyr: Tidy Messy Data*. R package version 1.3.1. https://CRAN.R-project.org/package=tidyr

Wilcox, R. R. (2022). Chapter 3 - Estimating measures of location and scale. In R. R. Wilcox (Ed.), *Introduction to Robust Estimation and Hypothesis Testing (Fifth Edition)* (pp. 45–106). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-820098-8.00009-9

Wood, S. N. (2011). *mgcv: Mixed GAM Computation Vehicle with GCV/AIC/REML Smoothness Estimation*. R package version 1.9-1. https://CRAN.R-project.org/package=mgcv

Yanagida, T. (2024). *misty: Miscellaneous Functions*. R package version 0.6.5. https://CRAN.R-project.org/package=misty

Yang, E. (2017). Bilinguals' working memory (WM) advantage and their dual language practices. *Brain Sciences*, *7*(7), 86. https://doi.org/10.3390/brainsci7070086

Zhou, B., & Krott, A. (2016). Data trimming procedure can eliminate bilingual cognitive advantage. *Psychonomic Bulletin & Review*, *23*(4), 1221–1230. https://doi.org/10.3758/s13423-015-0981-6

Appendix

Overview of Variable Codes

Measure	Variable code	Description
Demographics	participant_id	Subject identifier
	group-code	Group identifier for general (G) population, polyglots (P), and hyperpolyglots (H)
	Sex	Biological sex (M–F)
	age	Biological age
	edu	Years of education
	nlang	Number of spoken languages
Entropy	entropy_competence_speak	Entropy for all spoken languages
CVLT	CVLT_tot-imm	Sum of correctly recalled items in part 1 (short + immediate)
	cvlt-long_RT	Average reaction time per trial for recognition task
	CVLT_short	Correctly recalled items in short term recall of list A
Digit span task	digit-for_span	Maximal span length forward
	digit-back_span	Maximal span length backward
CBT	corsi-for_span	Maximal span length forward
	corsi-back_span	Maximal span length backward
NWRT	non-word-rep_span	Maximal span length
	non-word-rep_acc	Accuracy of correct items
Spoonerisms	spoon_acc	Sum of correct items Spoonerism task
	spoon_RT_manual	Total task duration in sec Spoonerism task
AMMA	amma_tonal	Score on tonal change identification
	amma_rhythm	Score on rhythm change identification
	amma_total	Total score performance on tonal and rhythm change identification