UNIVERSITÉ DU LUXEMBOURG

PhD-FHSE-2024-28
The Faculty of Humanities, Education and Social Sciences

DISSERTATION

Defence held on 02/07/2024 in Esch-surAlzette

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN *PSYCHOLOGIE*

by

Rémy Fabiano Lachelin
Born 15 September 1991 in Geneva, (Switzerland)

BILINGUAL LEXICAL AND SEMANTIC REPRESENTATION OF NUMBERS

Dissertation defence committee

Dr Christine Schiltz, dissertation supervisor
Professor, Université du Luxembourg

Dr Elise Klein
Assoc. Professor, Université Paris Cité

Dr Antoine Fischbach, Chairman
Assoc. Professor, Université du Luxembourg

Dr Amandine Van Rinsveld
Université Libre de Bruxelles

Dr Sonja Ugen, Vice Chairman
Université du Luxembourg

“El Dabih non mostrò alcun stupore. Prese una penna, scrisse alcuni numeri su un foglio. “Grande Scorpione,” disse alla fine, mostrandoglielo, “noi arabi inventammo questi numeri: il sistema decimale. Ma la nostra più grande invenzione fu syfr. Syfr, che divenne poi zephirus e poi zero. Noi inventammo il numero che indica il vuoto, il nulla. Un numero pauroso, nel cui segno circolare la mente si può smarrire.

Ebbene, tu conosci lo zero, esso è il numero delle grandi cifre; aggiunto, in lunga fila dietro a un semplice numero, lo trasforma in un mostro: un miliardo, un miliardo di miliardi. Sono i numeri con cui si indicano le tue grandi ricchezze: e lo zero vi cammina in fila, come in una carovana i cammelli carichi di gemme e sete, dietro al padrone. Esso è il tuo servo fedele: uno zero. Il tuo popolo, tanti zeri dietro a te, e così i tuoi consiglieri. Io potrei forse essere il secondo o terzo zero, nel grande numero della tua gloria: ma sempre vuoto, uguale a tutti gli altri. Ma non è questa la sola cosa che ti sfugge.

Lo zero spalancò anche un'altra via: se lo zero si fa seguire da una virgola, e poi da altri numeri, ebbene non ci sarà numero, per grande e mostruoso che sia, che potrà uscire dal suo orizzonte. Esso crescerà, schiererà cifre come soldati, ma sarà sempre, ahimè, meno del numero più piccolo, meno di uno. Così tu rincorri un potere assoluto, ma per quante cifre, numeri e soldati vivi e morti tu possa mettere insieme, davanti a te c'è uno zero: il mistero che non afferri, la natura, che supera ogni tua ricchezza, il cielo, che non puoi avvicinare. E bada! Dopo lo zero, e la virgola, possono seguire molti, altri zeri. Milioni di zeri. Ma se alla fine ci sarà un numero, esso esisterà. Questo è il mondo che non ti appartiene, la via che ti sfugge, l'infinitamente piccolo della libertà nascosta, il mistero della complessità che non puoi avere.

“El Dabih showed no astonishment. He took out a pen and wrote some numbers on a sheet of paper. "Great Scorpion," he said at last, showing it to him, "we Arabs invented these numbers: the decimal system. But our greatest invention was syfr. Syfr, which later became zephirus and then zero. We invented the number that indicates emptiness, and nothingness. A frightening number, in whose circular sign the mind can get lost.

Well, you know, zero is the digit of the great numbers; added, in a long line behind a simple digit, it turns it into a monster: a billion, a billion billion. These are the numbers by which you denote your great richenss: and the zero walks in a row behind it, as in a caravan the camels carrying gems and thirst, walk behind their master. It is your faithful servant: a zero. Your people, so many zeros behind you, and so are your councillors. I could perhaps be the second or third zero, in the great number of your glory: but always empty, equal to all the others. But that is not the only thing you are missing.

The zero also opened up another road: if the zero is followed by a comma, and then by other digits, well, there is no number, however great and monstrous it may be, that can escape its horizon. It will grow, and deploy digits like soldiers, but it will always, be less than the smallest of numbers, less than one. So you will chase after absolute power, but no matter how many digits, numbers and soldiers living and dead you may put together, there is a zero before you: the mystery which you cannot grasp, the nature, which surpasses all your richness, the sky, which you cannot approach. And beware! After the zero, and the comma, many, many more zeros can follow. Millions of zeros. But if there is a number at the end, it will exist. This world does not belong to you, it escapes you, the infinitely small of hidden freedoms, the mystery of complexity which you cannot possess.

"Il tuo più povero suddito è un numero, in fondo a tanti zeri: ma esiste, è vivo. C'è chi ammira le grandi misure e i grandi numeri necessari per esprimere la grandezza dell'universo, le distanze delle stelle. Ma lo scienziato, e l'uomo comune, resterà parimenti stordito dai numeri che inseguono e trovano la più piccola particella atomica, l'occhio dell'ape, la cellula.

Questa vita che hai intorno, i tuoi sudditi, la natura, ciò che sta nell'altra terra lontana dello zero, tu la disprezzi. Vorresti cancellarla. Pensi che tutto si possa comprare, pensi che i tuoi numeri siano abbastanza grandi per abbracciare il mondo. Essi sono syfr, zephir, il nulla, il vuoto. Le cose che tu puoi comprare sono un numero così infinitamente piccolo, che dovresti vergognartene. Non gloriarti della tua ricchezza. Essa è niente, sia se la rivolgi verso il cielo, sia verso i mondi dell'infinitamente piccolo."

"Your poorest subject is a number, at the bottom of many zeros: but it exists, he is alive. Some admire the great measures and large numbers needed to express the size of the universe and the distances of the stars. But the scientist, and the common human, will equally be stunned by the numbers that chase and find the smallest atomic particle, the bee's eye, the cell.

This life you have around you, your subjects, nature, what is in the other land far from zero, you despise it. You would like to erase it. You think everything can be bought; you think that your numbers are big enough to embrace the world. They are syfr, zephir, nothingness, emptiness. The things you can buy are such an infinitely small number that you should be ashamed of them. Do not glory in your wealth. It is nothing, whether you turn it towards the heavens or the worlds of the infinitely small."

Terra!, Stefano Benni, 1996, p. 161

English translation from the author

"Karma police, arrest this man, he talks in maths
He buzzes like a fridge, he's like a detuned radio"

Radiohead, Karma Police

Acknowledgement

Foremost, thanks to Christine for being the most amazing supervisor I could imagine! You have been amazing on so many different levels. For the deep contents of scientific discussions, availability despite an agenda so full it hurts your eyes just by looking and your open mindset. For your bulletproof positivity and pragmatic realism, which helped me see beyond many obstacles. For your deep sense of respect and humanism, permeates important decisions and is a huge inspiration for me and the team. For your heterogeneous horizons of interest spanning through music, cinema, theatre and dance making discussions beyond research questions as much interesting. Christine, from the bottom of my heart, thanks for all your precious help and support.

Thanks to Aliette, Cintia and Stella for being such amazing office mates, thanks for all the formal and informal discussion and the positivity I felt in the office. Dear team(s), thanks for the collaborative spirit and ambience. Thanks to the first team wave Alex, Max, Anna and Carrie. Thanks to the second team wave Tânia, Amaury, Hoyeon, Claire and Marion for the coffee breaks, crazy discussions and most importantly board games nights. It was nice to be surrounded by other like-minded PhDs. I want to thank Aliette, Talia and Mila for being such inspiring great scientists! Thanks to the “Louvain-la-Neuve delegation”: Margot, may the spirit of Beyonce stay alive in you, Cathy, I will always remember the intense testing session in LLN, it was always a pleasure to go to conferences knowing we would meet. Virginie thanks for your unforgettable generosity during my stay in Louvain-la-Neuve. Thanks also to the amazing and dedicated groundwork done in the lab by student assistants: Gina, Laura and Sarah!

Nicolas, you impressed me with your skills with the vending machine, so much free chocolate! Joke aside, I was very inspired by your sharp critical sense and sense of humour. Thanks also for your contagious passion for board games and all those movie recommendations. Mila, thanks for your determination, in insisting that covid is not a good pretext to lower our research goal. Without you, study 3 would not exist. Ilse, my “MCLS buddy”, thanks for our meeting discussions on the strangeness of working in research as a PhD candidate.

I moved to several shared apartments and lived with wonderful people: Konstantin, Laura, Annalea and Jenna in Trier. Amira and Kete in Göttingen. Amaury, Marie and Claire in

Audin-le-Tiche. Marc, I never thanked you for the room so close to the university which was a much-needed relief on my commuting routine at that time. Dodo and Virgile in Strasbourg for welcoming us in their shared apartment and being so adorable with Alba. Also unconventionally, I must thank Luxembourg for providing me with a generous parental leave granting me the resources to endure both parental and PhD endeavours.

Thanks also for the support of my best friends: Yveline and Marie, Naïma and Prisca, Fredi and Gaia and Feliça. And the deeply resourcing holidays and strong supporting love from the "butter tribe": Guillaume, Laura, Fabien, Mickäel (team 13 représente), Yann, Kevin and Timothé.

Mia and Alba, thanks for our family and all the love. We had a rough year, and I am astonished at how strong our love is keeping together, remember dancing in the kitchen. I am grateful for the help of Oma Ether and Nonna Eliane for your support and for having taken care of Alba while I was busy writing those lines and Mia studying.

Mia, thanks for your support (meant to be the French "de me supporter"), I am grateful for every step we walked together and the adventures that will come. Ti amo!

In memory of my brother Laurent Lachelin and my uncle Philippe Lachelin

Table of Contents

Table of Figures[1]

Note about the formatting: while the GENERAL INTRODUCTION and GENERAL DISCUSSION are formatted according to APA 7th guidelines, the three studies might differ in formatting according to the editorial line of the journal in which they were accepted/submitted. Hence for example Figure and Table captions might vary across this thesis. Also, Figure and Table numberings restart at 1 for each section. The terminology between GENERAL INTRODUCTION, GENERAL DISCUSSION and the three studies might also be inconsistently used.

---

[1] Only Figures of the GENERAL INTRODUCTION and DISCUSSION

Declaration of Authorship

I hereby declare that I am the sole author of the work entitled:

Bilingual lexical and semantic representations of numbers

and here enclosed, and that I have compiled it in my own words, that I have not used any other than the cited sources and aids, and that all parts of this work, which I have adopted from other sources, are acknowledged and designated as such. I also confirm that this work has not been submitted previously or elsewhere.

The 21st of May 2024

Date Signature

Rémy Fabiano Lachelin

First name and surname

General Introduction

1 GENERAL INTRODUCTION

Numerals are omnipresent in many parts of our modern daily life. Numerals let us measure and quantify physical properties of our realities such as space, time, or the effect of gravity as well as abstract concepts such as economic value with money. As the following example of a clafoutis recipe demonstrates, numerals are essential for communicating about quantities.

> *Warm up the oven to 210 C°. Wash and prepare 600 g. of cherries. In a bowl, mix 100 grams of flour, 60 grams of sugar, a little bit of vanilla and a pinch of salt. Then, add little-by-little 4 eggs, 2 dL of milk and 40g of melted butter, mixing in between. With some butter, grease a baking dish. Add the cherries and the preparation from the bowl to the dish. Place the dish inside the oven for 10 minutes and then lower the temperature to 180 C° for the last 20 minutes. Serve the clafoutis while still warm or cold with powdered sugar on top.*

This delicious example illustrates the importance of numerals to communicate and replicate a recipe. The numerals precisely indicate the quantities of each ingredient and the cooking temperature and time to successfully replicate the recipe. For example, imagine the beginning of this recipe but without numerals: "*warm up the oven to very hot. Wash and prepare a lot of cherries"*, this would inevitably lead to confusion and not make you succeed in backing a clafoutis. This is because numerals elicit an exact representation[2] of numbers in our minds. This representation can be manipulated mentally, for example after mixing 100 grams of flour and 60 grams of sugar, we can mentally calculate that the content of our bowl weighs exactly 160 grams. The recipe example also illustrates that without numerals it is also possible to verbally convey estimates of quantities, such as "a little bit of" or a "pinch of salt", which is fine for small but not large quantities. From this example, we can deduce why educating individuals about numeracy: using numerals, representing numbers and doing mathematics is so relevant for modern societies that compulsory education is imposed by most governments in the world.

[2] In this thesis, a mental representation is defined as pattern of brain activation that correspond to the external environment (see M. Johnson & Munakata, 2005). Hence numerals are the external environment's symbols (i.e. 5) while numbers are their mental representation.

Despite popular belief, numeracy and mathematical learning are not independent of languages. Indeed, not only does mathematical teaching occur using language and its specific vocabulary but also the verbal representations of numbers depend on languages. Hence, with over half of the global population being bilingual (Grosjean, 2010) the investigation of verbal representations of numbers in bilinguals becomes especially significant. This question is particularly critical for the multilingual education system of Luxembourg where the school curriculum changes from being taught in German to French. Luxembourg's multilingualism is represented in many languages: Luxembourgish, German and French and other languages such as Portuguese. Hence the Luxembourgish educational system faces a challenging difficulty: providing high education standards for literacy and numeracy for students with different language backgrounds on top of high proficiency in German and French. Hence the importance of the question raised in this thesis: how do bilinguals' lexical and semantic representation of numbers compare across languages?

In the following, I will begin by describing the approximate preverbal foundations of numerical representations (§ 1.1 Preverbal approximate number representations). These preverbal representations are approximate in the sense that it is not possible to do fine-grained distinctions of quantities (*i.e.* when the ratio between quantities is small). Symbolic numerals such as visual (*i.e.* Arabic numerals) and verbal (*i.e.* number words) sustain exact representations of numbers and are described in (§ 1.2 Symbols for exact number representation and § 1.3 Number words for exact numerical representation). I will then focus on the verbal aspect of number representations, shortly describing how counting is learned (§ 1.4 Acquisitions of exact number representation). Then, we will see several theoretical cognitive models for those different representations of numbers in the following chapter (§ 2 Models of numerical representations).

1.1 Preverbal approximate number representations

The **approximate** representation of numbers is sustained by a preverbal and evolutionary ancient cognitive system. This system is sometimes called the approximate number system (ANS) or the "number sense" (Dehaene, 2011). The ANS enables the estimation of physical quantities across modalities such as the quantity or numerousness of visual objects, sounds or tactile stimuli. It is called "approximate" since the discrimination between quantities is fuzzy. For example, the discrimination accuracy of the ANS depends on the ratio between quantities: 5 balls *vs* 15 balls (ratio 1:3) are easier to discriminate than 10 *vs* 15 balls (ratio of 2:3). More generally, the ANS follows the Weber-Fechner law: there is a logarithmic relation between the amount of stimulation to be noticed or discriminated from a reference stimulation.

The **ANS** is an evolutionary ancient cognitive system since its function can be found in other species. Animal comparative behavioural sciences have found that non-human animals such as non-human primates (Brannon & Terrace, 1998; Rumbaugh & Savage-Rumbaugh, 1987), rats (Meck et al., 1985), pigeons (Roberts & Mitchell, 1994), chickens (Rugani et al., 2008), salamander (Uller et al., 2003), frogs (Stancher et al., 2015) or even bees (Howard et al., 2018) are also able to do small calculations or discriminate between quantities within the Weber-Fechner range. The ANS is also found in babies and newborns in typically developing human beings. 6 months old can already discriminate between sets of elements, given the ratios are large enough such as 8 vs 16 or 8 vs 12 (Izard et al., 2009; F. Xu & Spelke, 2000). The ANS has an abstract function and works across modalities (i.e. visually, acoustically and cross-modally Barth et al., 2003).

There is an additional preverbal core cognitive system that qualitatively differs from the ANS, the **subitizing** system (Feigenson et al., 2004). Subitizing allows for a precise and fast perceptual apprehension of small quantities, i.e. 1 to ~4 (Kaufman et al., 1949). For example ●●● or ●●●● can be identified as 3 or 4 items very quickly (hence the term subitizing, from the Latin “subitus” immediately). This perceptual ability might be sustained by the Object Tracking System (OTS or Parallel Individuation): a cognitive system that allows one to represent individual objects in parallel until 4 or 5 (Carey et al., 2017; Hyde, 2011; vanMarle et al., 2018). This system is already found in very young babies (Starkey & Cooper, 1980).

Taken together, this suggests physical numerical quantity **approximate** discrimination might be a cross-species cognitive ability which is active very early in life or even innate. The limit of this pre-verbal system for quantities is that it is precise only until three or four elements (i.e. in the subitizing range) and becomes increasingly approximate – or fuzzy – with increasing quantities. In sum, typical human beings are born with an evolutionary ancient core cognitive system to approximately represent numbers, which is pre-verbal and hence independent of language (Gelman & Butterworth, 2005). This system is limited to 3 or 4 items (i.e. the subitizing range), above which its preciseness decreases leading to estimations. Hence counting is necessary for precise discrimination and exact representation of large numerosities.

1.2 Symbols for exact number representation

The greatest limit of the preverbal system is that it is only **exact** for small quantities under three or four. It is therefore not suited for larger exact quantities. This limit becomes relevant in environments where exact storage and retrieval of large quantities is important, such as modern societies. For example, when possessing a flock of sheep, being able to count them to know if and how many have been lost, is an important ability. The appearance of **symbols** allowed the exact representation of numbers. For instance, fingers on a hand can symbolize “five”, or written visual symbols such as V or 5 indicating an exact quantity. For example, discriminating 584’293’285 vs 584’293’281 dots is nearly impossible, while the same quantity represented by Arabic numerals allows a numerate individual to quickly find the biggest and even find the exact difference of 4 between both very large numerals representing an enormous numerousness. Furthermore, numerical symbols allow not only exact quantification but also persistence through time via memory or writing and exact calculation (i.e. 2 + 2 = 4, not 5). However, numerals (numerical symbols) are not intuitive as for the ANS and need to be learned. In turn, learning numbers seems to improve the ANS (Shusterman et al., 2016).

Epistemologically, it is conceivable that the initial method of symbolizing numerals might have been through fingers, which might explain the prevailing use of base-10 counting systems nowadays. The limit of fingers however is their quantity, i.e. 10 (or 20 including feet’s fingers). Hence a more complex counting system has developed using body parts, allowing one to count to 41 in the version presented in Ifrah & Bellos (2000). Body parts are however hard to store in the long term, hence visual symbols have largely replaced them as counting devices. Interestingly those symbols went from more analogic (such as I I I I I) to more abstract (V, in Roman numerals) and finally entirely symbolic such as 5. One of the advantages of Arabic numerals is that larger quantities can be stored using less space (i.e. MMXXIV = 2024) hence facilitating decoding and learning (Ifrah & Bellos, 2000). Other advantages of Arabic numerals, which have contributed to their success, include the use of the place-value system (i.e. the leftmost digit represents units, then tens, etc.) and the numeral 0, representing emptiness alone (i.e. 0) and as a place holder for multiples of tens (i.e. 1000). Visual symbols including Arabic numerals have been intensively taught to non-human animals such as chimpanzees. These experiences requiring many repetitions lead to a lot of errors and could interestingly be thought only until four (Boysen & Berntson, 1989). Similar experiments have been led with honeybees, leading them to successfully associate a symbol with a quantity above chance levels. However when presented with the reverse (quantity → symbol, i.e. ●●●●●● → 6) performances fell below chance level (Howard et al., 2019). Hence indicating that teaching visual numerical symbols to animals is limited to quantities in the subitizing range or by unidirectional associations.

For human beings living in modern digitalized societies **numeracy,** the mastery of numerals and their representations, plays a crucial role. For example for navigating in modern environments (i.e. bus, floors, or street numerals), orienting in time (i.e. seconds, hours and dates) or managing resources such as money (i.e. earning, buying, and selling). Numerals and their representations are also essential for measuring, comparing, and quantifying. For this reason numeracy, alike literacy, has become a fundamental pillar of education. Furthermore, the importance of numeracy is reflected in real life: mathematics level in school predicts later careers (Duncan et al., 2007), suggesting its importance for individuals. Since mathematics is a discipline taught in schools and hence relevant for practitioners in education, many studies in education have focused on how children and adults perform in mathematical problem-solving and arithmetic.

1.3 Number words for exact numerical representation

Besides visual symbols, an exact representation of numbers can also be sustained verbally with number words[3]. Since number words are related to oral traditions, they vary more than Arabic numerals: each languages have different number words (Comrie, 2013). For example in Western Europe: “eight”, “acht”, “huit” and “otto” vary across languages (despite likely sharing common etymological Indo-European roots) while 8 is commonly in use across the continent. Hence, in numerate individuals, the exact representation of numbers takes place in parallel with Arabic numerals and number words.

Number words are the lexical elements necessary for the exact verbal representation of numbers. This statement has been proven by several cross-cultural studies investigating populations using languages with restricted number word systems. For example, the Pirahã and Mundurukú languages do not have number words above five (i.e. a word that can be translated

[3] Terminologically I should call them numeral words, but I will remain with the most commmon terminolog “number words“.

as “many” would be used for bigger numerosities). When asked how many items they see, Pirahã and Mundurukú speakers reliably use the correct number words when the quantity indicated is below five. However for larger quantities the use of different words for the same quantity (hence the reliability) increases (Gordon, 2004; Pica et al., 2004). Yet, cross-linguistic differences are often confounded with cross-cultural differences. Thus the approximate representation of numbers in the above-mentioned populations could alternatively be caused by the fact that both societies are non-numerate. To circumvent this issue, a particular Nicaraguan deaf population has been investigated. Due to a lack of adapted education, Nicaraguan deaf were not educated to learn sign language. Hence, they are used to communicate with homemade gestures. Similarly to Pirahã and Mundurukú speakers, those homemade gestures lack for number words above five but Nicaraguan deaf live in numerate societies (i.e. they are familiar with the use of money). Similar approximate representations for numbers above three have been found in Nicaraguan deaf home signers (Spaepen et al., 2011). These studies therefore show that language provides access to exact numerical representations since restrictions in number word vocabularies drastically impact the preciseness of large number representations.

As an intermediate summary, we have seen that numerosities can be represented in three different formats or codes. Numerals are represented approximately by the ANS and exactly by two symbolic systems: visual symbols such as Arabic numerals and number words[4]. While the ANS is an intuitive system, the representation of symbolic numerals needs to be learned. In modern societies, the acquisition of numerical abilities takes place informally through parental and social education, and formally through obligatory school education.

1.4 Acquisitions of exact number representation

During the development of human numerate individuals, number words are acquired before Arabic numerals (Benoit et al., 2013; Le Corre & Carey, 2007). Arguably, the first step in the acquisition of number words involves **counting**. Counting is learned at an early stage and fosters the acquisition of numeracy concepts, contributing to building exact numerical representations. Counting relies on number words: a set of arbitrary words following a

[4] These three codes are represented in distinct parts of the brain as we will see in § 2.2 Triple Code Model (TCM)

conventional order (“one”, “two”, ”three”, etc.)[5]. In counting, the last number word designing the element of a set corresponds to the cardinality of the set: the total number of elements (i.e. ● “one”, ● “two”, ● ”three”, ● “four”, ● “five”, where *five* is the total number of dots). Therefore, learning to count co-occurs with the learning of number words, suggesting a substantial influence of language.

The question of the acquisition of precise numerical representations is still debated to this day (see for example Sella et al., 2021; Spelke, 2017), with two concurrent theories. In the first set of theories, the **ANS** plays a predominant role such that numbers are represented as an approximate set of objects (i.e. ●●●●● ~ iiiii). Hence number words are acquired by association or mapping number words with the ANS (i.e. ●●●●● – five, ●●●●●● – six, etc.), followed by the association of these number words with Arabic numerals (i.e. five - 5) (Benoit et al., 2013; Odic et al., 2015). Hence the meaning (or semantic) of number words as well as the ordinal nature of the counting system depends on this mapping (Lipton & Spelke, 2005). Traces of this mapping might be found in the observation that the ANS is involved in the very early preverbal ability to detect violations of simple additions such as 1 + 1 objects equal 2 objects (Wynn, 1992a). Further, sustain for this theory is found in studies finding support for the ANS for simple arithmetic operations (Feigenson et al., 2004; Geary et al., 2015). Or by the correlation between the ANS in children with their later mathematical performances (Starr et al., 2013). Or in that, the early acuity of the ANS predicts later mathematical performances (i.e. Halberda et al., 2008, 2012).

The second set of theories however, argues that the ANS might have a minimal or no role in early number acquisition and that initial number acquisition takes place with the support of the Object Tracking System (**OTS** or Parallel Individuation). For the OTS theory, numbers are represented individually and in parallel for items from 1 to 4 or 5, i.e. ●●● = i, j, k (Carey et al., 2017; Hyde, 2011; vanMarle et al., 2018). Hence for the OTS account the acquisition of symbolic numerals bypasses the ANS by constructing symbolic semantic associations among numbers (Carey, 2009). One of the criticisms of the ANS account described in the previous paragraphs is that the ANS does not provide the successor function which is necessary to learn

[5] Although fingers, body parts or an external tool such as small rocks could also be used to store the exact quantity.

to count (Carey, 2009; Carey & Barner, 2019). The successor function is defined in that each number is succeeded by another one (Carey, 2009; Wynn, 1990, 1992). Hence, the first developmental step would be the association of small sets of items with different verbal labels (i.e. ● – mama, ●● – papa, ●●● – sister, etc.), moving then to standard number word labels (i.e. ● – one, ●● – two, ●●● – three, etc.). In this set of theories, language plays a bigger role in numerical concept acquisition than the ANS account. This is underlined by studies relating vocabulary knowledge with early number knowledge (Negen & Sarnecka, 2012) and the finding that reading fluency is strongly predicted by counting skills (Koponen et al., 2013). Moreover, more recent findings did not find a correlation between performance with non-symbolic and symbolic numerals (Holloway & Ansari, 2009; Sasanguie et al., 2014)

Also, alternative explanations for the acquisition of symbolic numerals have been put forward in contrast to the ANS mapping account such as the bootstrapping method. Once the first four number words are acquired with the OTS, the second step would be to generalize the knowledge of these to all other number words. This step of generalizing the use of the first number words to larger numbers might use a bootstrapping method. Bootstrapping methods are based on inferring more complex rules from previous simple ones (Carey, 2009). This account is argued, would explain why the acquisition of cardinality is relatively slow: it takes about two years to understand that the last number words in the list corresponds to the cardinality of the set (i.e. cardinal principle knowers). Cardinal principle knowers have therefore implicitly acquired the successor function: that each number is succeeded by another one (Carey, 2009; Wynn, 1990, 1992).

2 Models of numerical representations

Several cognitive models have been designed for the representation of numbers. The following models differ in that they have originally been developed to describe how arithmetic is processed, or for how numbers are processed. Transcoding is the process by which numbers are converted into different codes such as:

- Analogic to visual such as dots naming or enumeration tasks: ●●●●● → /five/
- Visual to verbal such as number naming tasks: 5 → /five/
- Verbal to visual such as number dictation tasks: /five/ → 5

- Visual/verbal to analogic such as in number line estimation tasks: draw where /five/ or 5 lies on |1__________10|

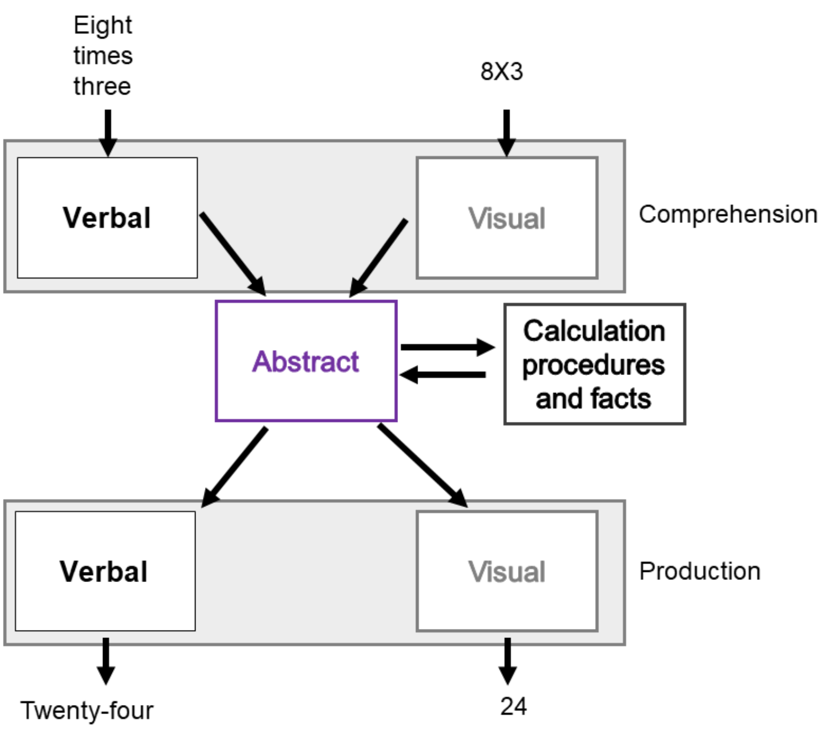
Another difference resides in their origins such as from observations, neuropsychological cases, developmental studies, etc. These models can be subdivided into **semantic** and **asemantic** (see Barrouillet et al., 2004), depending if the models postulate that all tasks involving numerals activate the semantic or not. Finally, more recent models are computational in that their predictions are based on mathematical formulae aiming to simulate human performances.

2.1 McCloskey's abstract code model

In their semantic model to account for number processing and calculation errors made by acalculic patients, McCloskey and colleagues (McCloskey, 1992; McCloskey et al., 1985). This model strong assumption is that there is an "abstract" module which is an obligatory passage for mental arithmetic. This model is subdivided into three main parts, see Figure 1.

Figure 1

Semantic model by McCloskey



Note: Adapted by the author.

First, for the comprehension of numerals, numbers are decomposed into lexical (elements) and syntax (relations) terms. Second, these terms were then passed to an abstract system which encoded the quantities such that ninety would have been: $90 = 9 * 10^1$, where 9 is the lexicon accompanied by the syntactic information that it is on the 1st power of 10, hence 90. Calculation procedures or long-term memory retrieval of the results are enacted by transforming the verbal or visual codes into abstract ones. Finally, in the third production part, verbal or visual codes ensured the composition of the results (output). Hence in this model, all numerical inputs are converted into an abstract, modality-independent, representation.

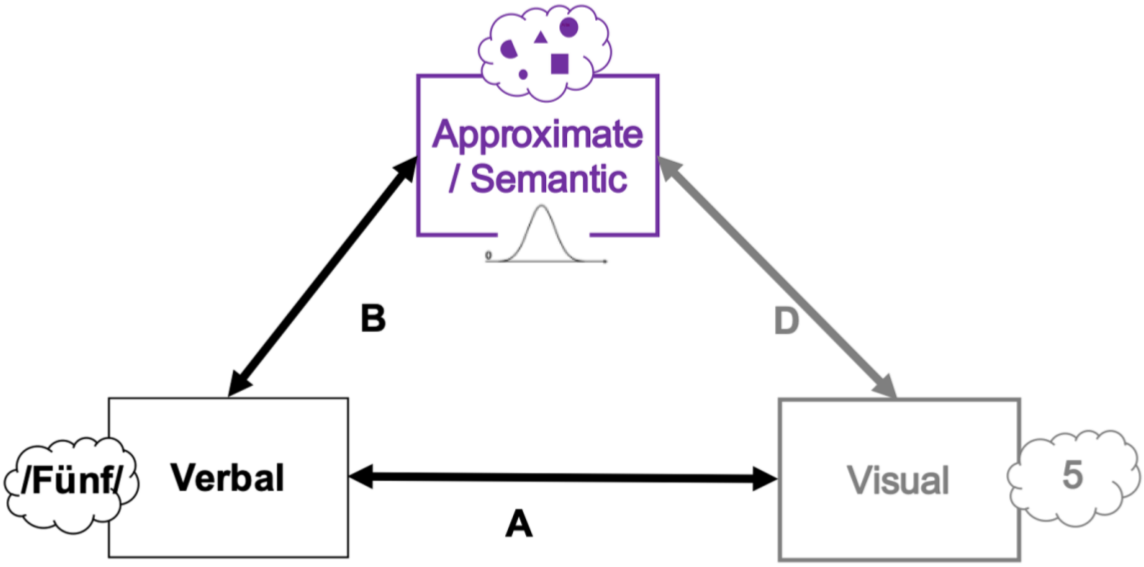
In a similar vein to McCloskey's model, Power and Dal Martello, (1990) who investigated children's errors, suggested a modification to the original model. In their modification, the semantic representation depends on the verbal structure of number words, for example, twenty-three thousand would be C1000 + ((C10*C2) + C3) (compared to $2 * 10^4 + 3 * 10^3$ in McCloskey's model). Hence for verbal to Arabic transcoding, the semantic representation depends on the verbal number word structure, rather than an abstract code.

2.2 Triple Code Model (TCM)

The triple code model (TCM) stipulates three codes for numbers which are represented in separate cognitive modules (see Fodor, 1985) and have different functions (Dehaene, 1992). The three codes are the approximate or semantic (*i.e.* ●●●●●), visual (*i.e.* 5) and verbal (*i.e.* "five") codes. The approximate code consists of the Approximate Number System (ANS) described in § 1.1 Preverbal approximate number representations. Hence it refers to a cognitive module found across species and hence being evolutionary ancient. On the other side, the verbal and visual codes are modality-dependent and culturally acquired. These three codes are interconnected by different routes, such that a visual nerals is bidirectionally associated with its semantic (line D) and verbal form (Line A). The verbal code is also connected with the semantic (B), see Figure 3.

Figure 2

Triple Code Model (TCM)



Note: Adapted by the author. Arrows: A indicates the association between verbal and visual codes of numbers. B indicates the association between the verbal and semantic code. C between the visual and semantic code.

The TCM is a **neurocognitive** model, in the sense that it posits functions and brain localization of the modules implementing those functions. The semantic representation is located bilaterally in the inferior parietal sulcus (IPS), the visual representation in the occipitotemporal areas and the verbal representation in the left Angular Gyrus (Dehaene et al., 2003; Siemann & Petermann, 2018). With regards to **verbal representations** of numbers which will be discussed in detail in this thesis, the TCM stipulates they depend on linguistic rules, which are sustained by neuropsychological cases (Delazer & Benke, 1997). For bilingualism, the TCM states that "*A strong ensuing prediction is that subjects must switch mentally between the two notations in the course of complex calculations. Such translation operations should introduce a measurable cost in RT.*" (Dehaene, 1992, p. 33, RT = Reaction Time). Hence implying that the verbal code is encoded in the language in which mathematics are learned. Therefore solving arithmetic in a second language is translated into the first, which would result in a cost, which is sustained by a large body of empirical evidence (see § 5.2 Language

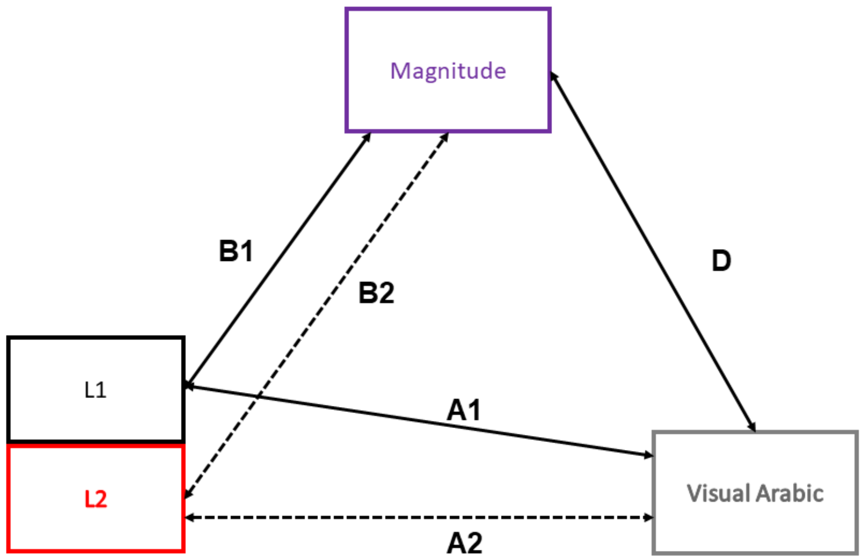
Switching Cost (LSC)). However, it also implies that the second language does not have a direct connection with the visual and semantic codes or representations.

2.3 Encoding Complex Model (ECM)

The ECM proposes the existence of format and modality-specific representations for each number code. Hence for arithmetic, the ECM suggests that arithmetic facts can be encoded for example in digit form (or Chinese characters as for Chinese speakers) or in number words. Regarding bilinguals, each language has independent representations of numerical facts (Campbell & Clark, 1988, 1992), see Figure 3.

Figure 3

Encoding Complex Model (ECM)



Note: Adapted by the author. A1 and A2 represent the association between verbal and visual codes of numbers. B1 and B2 are the associations between the two verbal codes and the magnitude code. D the association between magnitude and visual codes.

The ECM differs from the Triple Code Model that will be described later in that "[…] *communication between representational systems often involves interactive rather than strictly additive processes. Interactive processes are products of task-specific practice, which creates integrated encoding-retrieval processes within and between representational system*"

(Campbell & Epp, 2004, p. 231). Some early critics of the ECM include that the model is underspecified and that it is hard to use to make predictions (Dehaene et al., 1993).

2.4 Multidigit number writing model by Dotan and Friedmann

An additional model for number reading has been recently developed (i.e. Arabic to verbal number transcoding, i.e. 42 → "forty-two"), mainly from neuropsychological observation of different forms of acalculia (Dotan & Friedmann, 2018). This model presents two stages for number reading: first a visual analysis of the Arabic numerals and second a stage for the verbal production of number words. This separation between visual and verbal processes is directly inspired by the TCM (see also L. Cohen & Dehaene, 1991).

The **visual** analyser extracts the Arabic numeral's identity, decimal structure, and parses into triplets in a language-independent manner. For example, the decimal structure of 5840 is visually processed: it is detected as a 4-digit numeral and parsed into a triplet[6] ({5} and{8},{4},{0}), and the {0} position is detected as being the last digit. In parallel, each digit's identity and order is visually recognized to the 1 to 9 constituent digits in the correct order. Hence, at this stage, there have not been any lexical retrieval processes.

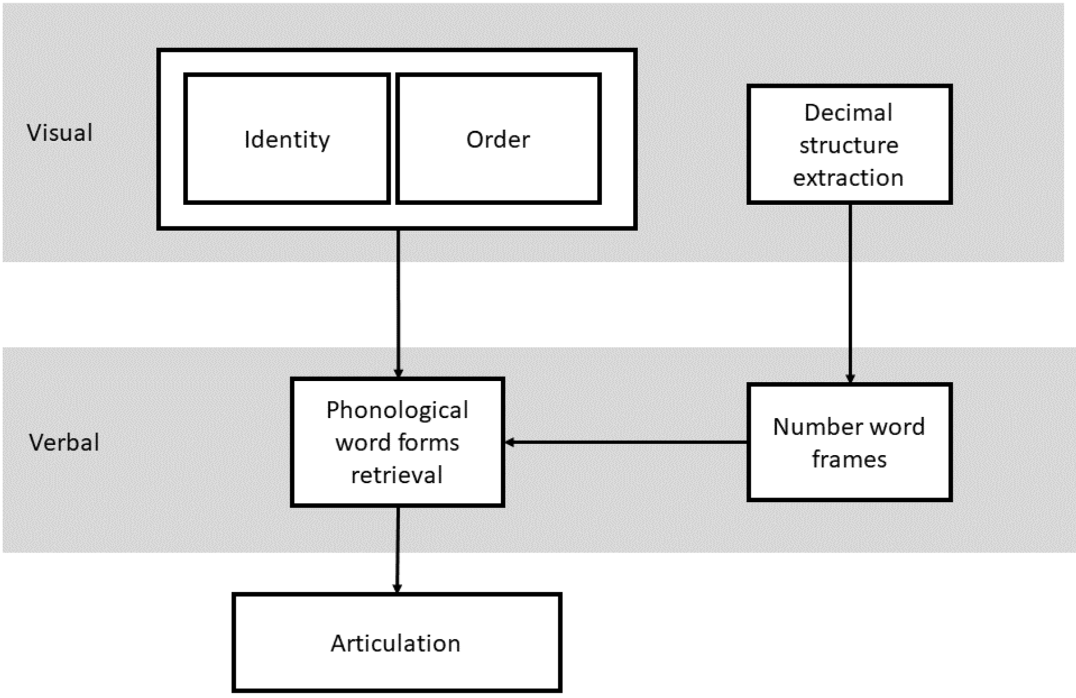
Then the information from the visual analysis of the decimal structure is passed to the **verbal** production system to linearly build a **number word frame** consisting of lexical class (i.e. owes, teens, tens), multiplier word (i.e. "hundred", "thousands") or the function word "and". For example 5840 becomes {_:ones} [thousand] {_:ones} [hundred] [and] {_:tens}. The number word frame is constructed in a hierarchical three like structure representation, analogous to the syntactic construction of sentences in linguistics. In parallel, the identity and order of each digit are visually passed to the verbal system which retrieves the phonological form of the number (i.e. 5 → /five/, 8 → /eight/). Hence the number word frame is fulfilled, so that for 5840 it becomes {5:ones} [thousand] {8:ones} [hundred] [and]{4:tens}. Finally, the word frame is assembled by morpho-phonological articulation procedures {/five/ /thousand/ /eight/ /hundred/ /and/ /forty/}. Hence, at this point only, language-specific rules are applied,

[6] The atuhor's note that the parsing might be language dependent such that in english numerals are organized into triplets, but for example in Chinese or Japanese these are organized in myriady (i.e. 4 digit chuncks). Hence influencing the intial construction of the syntactic tree.

such as the inversion in German. This assembly occurs in a phonological buffer from a dedicated phonological store which differs from other words.

Figure 4

Dotan and Friedman's (2018) multidigit number reading model



Note: Adapted by the author.

2.5 Developmental ADAPT model

Preceding number naming, (Deloche & Seron, 1982) proposed an asemantic model based on four processes: parsing, categorization of primitives, transcoding and production. Input numbers are first parsed and into lexical primitives matching with the number lexicon. Then the lexical primitives are categorized into a lexical class (i.e. units, teens or decades) and the position relative to this lexical class. For example 13 in English, is in the 3rd position of teens (i.e. eleven 1st, twelve 2nd and thirteen 3rd). Categorization also identifies hundreds, thousands or millions creating a three-slot frame. Finally, number writing is assured by the production process. However this model "*leaves little room for the learning of declarative knowledge*" (p.370 Barrouillet et al., 2004).

Hence later on a new model was proposed: A Developmental Asemantic Procedural Transcoding (ADAPT) model. Which is an asemantic model for number transcoding

(Barrouillet et al., 2004). It integrates a developmental perspective on how transcoding strategies change with age or experience. It is mainly developed to account for number writing, hence transcoding from verbal (i.e. /fünf/) to visual codes (i.e. 5). In ADAPT, children learning to transcode numerals use a procedural process which shifts to direct long-term memory retrieval in adulthood. The procedural processes include a parsing system which decomposes the input number word into parts. Then a set of procedural rules are applied to build the output Arabic numeral. These procedural processes are enacted by the working memory. In ADAPT increased frequency of retrieval of number words leads to more long-term memory direct retrieval by automatization of the process (Logan, 1988). With an increasing long-term memory retrieval strategy, the working memory can be used for other processes. Hence ADAPT is asemantic since the meaning of numbers is not involved in those processes. One of the predictions of ADAPT is that transcoding tasks in children working should involve working memory, but long-term memory for adults.

2.6 Computational Discrete Semantic System

In addition and sometimes complementary to the mainly modular models described above, there are computational connectivist models. These models are connectivist since they describe the quality of associations between different numerical representations. Hence, they are more suited for predicting the semantic effects of numbers. Importantly they are also computational, meaning that they can be resumed into mathematical equations with which it is possible to simulate performances and therefore directly test the predictions of these models.

Sparking from the debate of the influence of the ANS on the acquisition of numbers or not (see § 1.4 Acquisitions of exact number representation), an alternative model of abstract representation has been suggested: the discrete semantic system (DSS, Krajcsi et al., 2016; Sella et al., 2021). This model is discrete in that it describes the associations between symbolic representations (i.e. visual and verbal) without the involvement of the ANS. In the DSS the associations between the representation of symbolic numerals are determined by the same mechanism as lexical and semantic associations for words. Hence it is a connectionist model where the nodes are represented by the numbers (i.e. lexical representations) and connections between nodes form a semantic network. The distance effect would emerge from the strength of associations between numbers (i.e. closer numbers such as 4 and 5 are semantically more related than more distant ones such as 2 and 5). The size effect would emerge from the frequency of language occurrence of each numeral (i.e. smaller numerals are more frequent than larger ones (Dehaene & Mehler, 1992). The DSS and the ANS predictions concerning the number distance and size effects are slightly different (see Krajcsi et al., 2018).

2.7 Model's implication for verbal representations of numbers

The cognitive models for numerical cognition reviewed above commonly postulate a form of lexical, visual, and abstract or semantic representations for numbers. They differ regarding the role of the abstract or semantic representations, with the **semantic** models (i.e. McCloskey, 1992; Power & Dal Martello, 1990) postulating obligatory semantic processing (such as magnitude) of numbers when passing from a form of representation to the other (i.e. transcoding). While for **asemantic** models transcoding between verbal and visual codes does not require a semantic activation (i.e. TCM from Dehaene, 1992, Deloche and Seron (1982) model, Dotan and Friedmann (2018) model and ADAPT from Barrouillet et al., (2004) model).

Regarding the **lexical level**, all the models suggest an influence of language. However, the models differ regarding the lexico-semantic representation of number words. While the semantic model of McCloskey suggests an obligatory pre-activation of the meaning of numbers, later models do not imply it. The TCM suggests linguistic effects should be found when the verbal code is activated but not in the case of sole activation and interactions between the ANS and visual codes. The effect of language on verbal representations should be predictable by the same rules as language in general. However, the TCM makes no direct prediction for passing between verbal and visual codes. Finally, the ADAPT model suggests that linguistic features are particularly influential in the early learning of number words but that older adults activate them automatically by long-term memory retrieval, hence bypassing language-specific morpho-syntactic influences. Hence ADAPT should predict less efficient transcoding in less transparent languages which require more procedural rules to transcode in children but not adults.

3 Language-dependent influences on the exact verbal representation of numbers

In the models described above we have seen how languages can influence the verbal representations of numbers at **different processing levels**. Overall, all models suggest an effect of language either on lexical representation or on lexical retrieval or verbal production. Regarding language-dependent morpho-syntactic differences, these effects are explicitly described in Power and Dal Martello's, Dotan and Freidman and ADAPT model, mainly suggesting they arise from procedural mechanisms at the stage of verbal production. Then, regarding the meaning or semantics of numbers, McCloskey posits an obligatory semantic bottleneck (hence direct activation) when transcoding numbers, supposedly in all languages.

Additional levels of linguistic influences on number processing than the lexical, morpho-syntactic and lexico-semantic which are the focus of this thesis. For example, Bahnmueller and colleagues (2018), as well as Dowker and Nuerk (2016) identified six linguistic levels that might influence number processing: syntactic, lexical, visuo-spatial-orthographic, semantic, conceptual, and phonological. For the syntactic level it is intended the effect of language general grammar markings for quantities such as plural markers (see § 3.1 Language general: Grammar). The lexical level includes the effects of morpho-syntactic differences in the structure of number words. For example, concerning the base, the effect of inverted ten-unit place value in German number words: "zwei und *vierzig*" = *4*2, literally "two and *forty*" (see 3.2.2 Transparency of order) and the base-20 in "*quatre-vingt*-deux" = *8*2 in French, literally "*four-twenty*-two" (see § 3.2.1 Transparency of power). Those language-dependent differences can occur between languages and within languages, for example, some teen number words in English such as "thir*teen*" which is in a unit-*ten* order compared to *ten*-unit order for 20's to 90's numerals. The visuo-spatial-orthographic level regards the reading direction (i.e. comparing readers with left-to-right and right-to-left systems). The semantic level that is referred to in (Bahnmueller et al., 2018) concerns the use of quantifiers such as "much" or "many". It, therefore, differs from the definition of semantics that will be used in this thesis which is the qualitative association between numerals. The conceptual level is the level which associates certain numbers with certain concepts "*e.g., there are unmarked (even, right) and marked forms (odd, left) of most adjective pairs*" (Bahnmueller et al., 2018, p. 2). The phonological level regards the use of verbal working memory to process verbal numerals.

Another more overarching differentiation of the influence of language on numerical cognition is to subdivide them into language general effects and mathematical language-specific effects including number words. For language general, it is the influence of general language proficiency or exposure as well as linguistic characteristics such as syntax and grammar. For example, richer language exposure predicts faster number-word acquisition (Gunderson & Levine, 2011; Piantadosi et al., 2014) and more advanced skills such as general mathematical abilities (e.g., fractions and geometry; Kleemans et al., 2018; Kleemans & Segers, 2020; Vukovic & Lesaux, 2013).

3.1 Language general: Grammar plural markers

Languages differ in their grammatical structures, for example, to mark plurality or singularity, which also impacts grammatical differences in number naming systems and influences number processing. For example, comparing children speaking Russian and English that marks the difference between **singularity** and **plurality** (*e.g.* my *child is* waiting in the car) to Japanese that doesn't (*e.g.* my *children are* waiting in the car) shows that the learning of the meaning of one is delayed in the latter (Le Corre et al., 2016; B. Sarnecka et al., 2007). A similar pattern is found comparing Saudi Arabic and Slovenian to English which differ in the marking of singular, dual and plurality in the grammar and lead to earlier mastery of the concept "two" than in English (Almoammer et al., 2013).

3.2 Language specific morpho-syntactic transparency

Transparency is the degree of morpho-syntactic correspondence between visual and verbal number words, which consequently can vary across languages (Xenidou-Dervou et al., 2023). The degree of transparency hence depends on how many morpho-syntactic rules need to be applied for constructing number words compared to Arabic numerals. Arabic numerals are constructed on a positional value base-10 system. Hence the relative position of the Arabic numerals determines the value of its power. This can tentatively be summarized with $a * 10^{b-1}$, where a is the cardinal, 10 is the base-10 and b is the position of the numeral from left to right. For the example of 42, 4 is in the 2nd position hence $4 * 10^{2-1} = 40$ and 2 is in the 1st position so it is $2 * 10^{1-1} = 2$). Consequently, a change in the cardinal of the base occurs for each 10th numeral. For each *b+1* of 10^b an additional cardinal numeral is added on the left. For example, 842 becomes $8 * 10^{3-1} + 4 * 10^{2-1} + 2 * 10^{1-1}$. Regarding the morpho-syntactic correspondence with number words, a is the cardinal morpheme (i.e. one to nine) and b is the base position (i.e. base-10 in English). For example in "four-ten-two", "four" is a, "ten" is the base 10 and "two" is *b.* This would correspond to a highly transparent language for example Chinese as we will see in the following sub-chapter. For English, we can already notice an irregularity in "forty-two": "forty" is more opaque than "four-ten" and adds an irregular morpheme "for-" instead of "four" and "-ty" is also an opaque irregular morpheme replacing "ten". As we will see this opacity affects the processing of numbers (see § 3.2.1 Transparency of power). These examples demonstrate how morpheme's correspondence with number words can differ, leading to different degrees of transparencies. The syntax however remains intact, both in terms of the place values system and the base (i.e. $a * 10^{b-1}$). This is however not the case of English teen numerals (i.e. 11 to 19). First, "eleven" and "twelve" can be considered as additional number words whose structure is arguably too opaque to be decomposed morpho-syntactically (that would not be the case would they be "one-ten" and "two-ten"). Second, in teen numerals "thirteen" to "nineteen", there is both an opaque morpheme "teen" and a morpho-syntactic inversion of the place value. The term "teen" is opaque to "ten", the base (which can lead to the frequent confusion between for example "nineteen" and "ninety"). And finally, a syntactic inversion occurs between ten and unit, since 19 ($1 * 10^{2-1} + 9 * 10^{1-1}$) should be "ten-nine" to be morpho-syntactically transparent, and not "nine-te(e)n" (i.e. $9 * 10^{1-1} + 1 * 10^{2-1}$). Teen numerals have been proven to be more difficult to process in English (i.e. Ho & Fuson, 1998; Miller et al., 1995). Morpho-syntactic opacity occurs for most two-digit numerals in the two languages which will be compared in this thesis: German and French.

In **German**'s two-digit numerals, the ten and unit position is inverted. Hence 42 becomes "Zwei-und-Vvierzig", literally "two and forty" ($i.e.\ 2 * 10^{1-1} + 4 * 10^{2-1}$, instead of $4 * 10^{2-1} + 2 * 10^{1-1}$). In this case, there is a syntactic opacity with the Arabic numerals' ten-unit positional system. In addition to morpheme irregularities (i.e. "-zig" instead of "zen", the number word for ten). Note that numerals above two digits such as hundreds are ordered on the leftmost. Hence 842 is "Acht-hundert-zwei-und-vierzig", literally "eight-hundred-two-and forty" ($i.e.\ 8 * 10^{3-1}\ 2 * 10^{1-1} + 4 * 10^{2-1}$).

France's **French** number words have a mixed base-10 and base-20 system. While until the '60s number words are in base-10, number words between 70 and the base changes to a base-20 system. Hence 85 is "Quate-vingt-cinq", literally "four-twenty-five" instead of "eight-ten-five" ($i.e.\ 4 * 20^{2-1} + 5 * 20^{1-1}$, instead of $8 * 10^{2-1} + 5 * 10^{1-1}$). In this case, the syntactic opacity occurs in between the base-10 of Arabic numerals and base-20 number words.

French and German number words transparency of **powers** and **order** are therefore affected (see Bahnmueller et al., 2018 for a nomenclature). The transparency of power hence concerns the power of the base. Transparency of order regards the order of the positional value system, which is ten units (i.e. 42) in the Arabic number system.

3.2.1 Transparency of power

Several **Asiatic languages** such as Chinese or Korean are remarkably transparent. For example, the Chinese equivalent of 42 (forty-two), is literally "four-ten-two". Hence the morpho-syntactic structure of Arabic numerals corresponds to those of number words

($a * 10^{b-1}$), leading to being highly transparent. Learning such a number word system has several advantages for example in comparison to English: (1) fewer lexical units need to be learned (i.e. 10 lexical units, one to ten, allow to count up to 99) (2) there are no inverted teens (such as "nineteen") (3) there are no irregular morphemes (such as "forty") (4) the morpho-syntax matches Arabic numerals without exceptions (such as with teen numerals in English or the transparency of power and order of German and French seen above)[7]. Compared to for example English speaking American peers, the transparency of several Asiatic languages gives their speakers an advantage with regards to understanding the place-value system or Arabic numerals (Miura et al., 1993), number line estimation (Siegler & Mu, 2008) and arithmetic problem solving (McClung & Arya, 2018; Rodic et al., 2015). However, comparing different groups confound languages and cultural differences. For example, part of those differences can also be explained by educational and cultural differences such as time spent teaching mathematics or hours of homework and parental expectations (Stevenson et al., 1990).

To disentangle linguistics from cultural and social factors, Dowker and colleagues (2008) compared children in Wales who spoke English or Welsh, Welsh number words are more transparent than in English, such that eleven is "un deg un", literally "one ten one". In several reading and comparison tasks, the Welsh speakers outperformed the English monolinguals. Furthermore, Dowker and Roberts (2015) compared children English and Welsh speakers from the United Kingdom with a number-line estimation task. In this study, the participant's task was to draw where a number should lie on a line, (i.e. draw with an x, where 15 lies on this line: 0|____________|20). The results of the study show that the Welsh children were more accurate in their estimation than their English-speaking peers. Hence suggesting a direct influence of the language on a task that involves estimating quantities. Language transparency of power can also affect the numeral used as a base (i.e. the Arabic number system uses a base-10 system).

[7] Another advantage is that, most Chinese number words are monosyllabic, which could bring an advantage, since shorter words are easier to store in the verbal working memory (A. D. Baddeley et al., 1975).

Some languages use a **base-20** instead of a base-10 number system, such as Basque, Diola-Fogny or France's French (see Haspelmath et al., 2005). In base-10 systems, the change in base occurs each 10th, while in base-20 each 20th (i.e $a * 20^{b-1}$). French base-20 vigesimal system is however not uniformly adopted by all francophones, for example, in French-speaking Belgians and some Swiss-French regions, 90 is said "nonante" literally "ninety", and 70 is "septante" literally seventy (and in varying Swiss cantons as in Fribourg 80 is "huitante", literally eighty). The opacity between the base-20 structures of French number words compared to the base-10 structure of Arabic numerals leads to **increased difficulties**. Comparing the errors of French and Belgian French-speaking children in a number dictation task, (Seron & Fayol, 1994) found that French-speaking children made more errors in the 70's to 90's numerals. Functional error analyses of these errors revealed the syntactic errors were induced by the language structure of those numerals (i.e. 97 → 42017, literally "four-twenty and seventeen"). Saad (2010) investigated 6-year-old French-speaking children in a number dictation task and found a significant increase in the quantity of errors for numerals between 70 and 99. Camos, (2008) investigated 7-year-old children who also made more errors for vigesimal numerals compared to other numerals in a number dictation task. In a cross-linguistic study, Van Rinsveld and Schiltz (2016), compared English and French speaking 10-years-old children with a number reading (i.e. reading Arabic numerals) and an auditory-visual number recognition task (i.e. a numeral is heard and it must be matched with an Arabic numeral target among 3 distractors). Compared to English speakers, the French-speaking children had worse performances with base-20 numerals in both tasks. Similar results are also found in other languages using a base-20 number system such as in Basque. Colomé et al., (2010) compared adult Italian and Basque-speaking adults with addition problems and a number comparison task. Basque, like French has a base-20 rule for number words. Basque-speakers solved additions faster when they followed a base-20 structure, such as for 20 + 15 compared to 25+10, since the correct answer 35 is "hogeita-hamabost", literally "twenty-fifteen". Hence those results indicate that language influences the acquisition of numbers and also the solving of arithmetic problems, in that opacity with the Arabic numeral system leads to difficulties. Those difficulties might hamper or delay numerical acquisition, similarly as we saw for transparent Asiatic languages in comparison to less transparent languages such as English.

3.2.2 Transparency of order

Languages affected by the transparency of order invert the number word structure compared to the ten-unit positional system of for example Arabic numerals. Hence transparency of order mostly regards two-digit numerals. Arabic, Dutch, and German are some examples of inverted languages. The inversion hence impacts the transparency between number words and Arabic numerals and affects various tasks ranging from arithmetic to more basic numerical tasks. Note that the following studies concern the effect of two-digit and not multi-digit numerals(see Klein et al., 2013 for a review).

For **arithmetic**, the inversion can complicate the resolution of problems needing to carry units across decades. For example, when two numbers imply that adding the units leads to a decade change, *i.e.* 26 + 37 = 63. Problems with carry for 8 years-old German-speaking children have a larger cost in terms of reaction times, in comparison to their Italian-speaking peers (Göbel, Moeller, et al., 2014). Furthermore, German adult speakers show a cost for carry problems compared to Chinese speakers (Lonnemann & Yan, 2015). In a simple arithmetic task, 5-year-old Dutch-speaking children showed a cost related to *transparency of order* compared to their English-speaking peers (Xenidou-Dervou et al., 2015). On the other hand, the *transparency of order* might also be beneficial for example in terms of errors when verifying multiplication problems with common decades in the multipliers as in the solution (Bahnmueller et al., 2020). Arithmetic performances are longitudinally predicted by transcoding inversion errors and more compatibility effect (Moeller, Pixner, et al., 2011). Hence indicating that the inversion property plays a role in more basic numerical tasks.

Magnitude comparison tasks consist of judging which one is the biggest between two numbers. A compatibility effect is when a two-digit number magnitude comparison is easier to do when both ten and units are bigger (i.e. 45 < 78, 4 < 7 and 5 < 8) than when the unit is incongruent (i.e. 45 < 71, 4 < 7, but 5 > 1). A compatibility effect, i.e. slower responses for incongruent trials is consistently found in adults (Bahnmueller et al., 2015; Nuerk et al., 2005; Nuerk et al., 2001, 2004). To investigate the effect of reading direction on the inversion effect a magnitude comparison task with adult German, Hebrew, Arabic and English speakers was done. Hebrew and Arabic are written from right to left; however Hebrew number words are in ten-unit order (hence inverted relative to the Arabic numerals). It was found German and Hebrew readers had the strongest unit interference, interpreted to be due to the incongruence between reading direction and ten-unit order in number words, which was not observable for English and Arabic readers (Moeller, Shaki, et al., 2015).

In a **number dictation task** in first-grade German speakers, about 50 % of the errors could be explained by inversion (Zuber et al., 2009). Furthermore comparing first-grade French and Dutch speakers with a number dictation task, showed that Dutch speakers committed more inversion errors than French speakers also at second grades (Imbo et al., 2014), but these differences might have been explained by curricular differences (Krinzinger et al., 2011). The confounders given by curricular differences can be isolated using within-subject designs. For example, children speak Czech, a language containing two number word systems with opposed order transparencies, more error inversions occur for dictation tasks with the inverted compared to the non-inverted number words (Pixner et al., 2011). A study with 8-year-old German speakers (Steiner, Finke, et al., 2021) found more inversion-related errors in writing and reading numbers than in English-speaking children. Clayton and colleagues (2020) compared 6 and 7-year-old English and German-speaking children who heard a number in their native language and had to write it down in Arabic numerals. German-speaking children did more inversion errors for numerals above 20 than English-speaking peers and the performances of both groups correlated with the arithmetic task, suggesting common cognitive mechanisms underlying both. Interestingly for the following teen numerals 11, 13, and 16, which in English do not follow the conventional morpho-syntactic (i.e. “eleven” instead of “ten-one”) and are even inverted (thirteen instead of “te(e)n-three” and “sixteen” instead of “te(e)n-six”) the quantity of inversion errors did not differ with German-speakers. Another study found more inversion-related errors in a writing task made by German than Japanese-speaking children (Moeller, Zuber, et al., 2015). van der Ven et al. (2017) in a sample of about 25000 thousand children found that Dutch-speaking children make at least one inversion error when doing a transcoding web-based game.

In **auditory-visual number matching** tasks, the participant’s task is to match numerals in different formats such as auditory and visual. Auditory-visual number-matching tasks are interesting to investigate because their reaction times correlate with math performances, suggesting common underlying cognitive processes (Sasanguie & Reynvoet, 2014). To test the effect of inversion the auditory-visual matching tasks can be adapted so that after hearing the numeral 42 there could be an inverted distractor (i.e. 24). Children and adult German speakers are slower at rejecting inverted distractors than English speakers (Steiner, Banfi, et al., 2021). Another possibility is to precede a matching pair with the unit (i.e. _2) or the ten (i.e. 4_) information. This adaptation has led to faster reaction times in the ten first conditions for French (non-inverted) than German speakers (Poncin et al., 2019).

More dramatically, **neuropsychological** reports on adult German-speaking patients who, following an aphasic episode did not invert the two-digit numerals when reading them (Blanken et al., 1997). Another German-speaking patient, following a parietal lesion, has been found to write two-digit numerals from right to left starting with the unit and ending with the ten. This had been identified as a compensation strategy for his lesion, when asked not to use this strategy the patients would make many inversion errors (Sittig, 1921). A similar neuropsychological case is reported by Lochy et al., (2004): inversion errors in number writing were reported in a German speaker following a left temporoparietal infract. For example when asked to write down “Sieben-und-fünfzig” she would write 75 instead of 57.

To sum up, *transparency of order* is the inversion of two-digit numerals between units and decimals in comparison to their written visual form. Research with different inverted languages, among them German, shows a cost related to the inversion properties at different ages and for certain transcoding tasks which predicts later arithmetic performances.

3.3 Number word’s semantic

While the influence of morpho-syntactic transparency concerns the access to lexical number representation, we will see in the following how languages influence the access to the semantic representation of numbers. Semantic representation of numbers is defined herein as the qualitative associations between number representations. In the following, we will see the importance of language in learning to count and cardinality. The acquisition of numbers semantics is a crucial aspect of numerical development. Unlike regular words, number words derive their semantic meaning from mathematics, ensuring an exact semantic meaning. Number word learning is one of the building blocks on top of which mathematical competencies are acquired. Given that mathematics is learned cumulatively, meaning each previous acquisition step is required before the following one (Geary et al., 2013; Watts et al., 2014), counting, leads to access to arithmetic, geometry, functions and so on. For example how good children are at counting and doing basic numeric operations is a predictor of how they later perform in arithmetic (Krajewski & Schneider, 2009). Hence it is important to understand how languages influence learning to count and the acquisition of cardinality.

3.3.1 Cardinality and Counting

Before the counting and **cardinality principle** is acquired children learn to apply the one-to-one correspondence principle between number words and objects (see §1.4 Acquisitions of exact number representation). This principle consists of the coordination of partitioning sets into elements and tagging them (Gelman & Gallistel, 1978). Then the stable-order principle needs to be acquired, meaning the understanding that the number word list is stable since counting relies on the stable successive order of arbitrary number words[8] (“one”, then “two”, then ”three”, etc.). Finally, children learn to count: that the last number word matching the last element of a set corresponds to the total numerosity of elements: the cardinal (Gelman & Gallistel, 1978). For example when counting: “●●●●●”, “five” corresponds to the cardinal 5 of the set: hence it contains 5 x “●”. The acquisition of the cardinality principle is also sustained by the successor function: each natural number n has a successor $n+1$. It takes on average about two years for a child to acquire the cardinality principle, i.e. being a “cardinal principle knower” (Carey, 2009; Le Corre & Carey, 2007; Wynn, 1990, 1992a). This slow acquisition is argued to reflect an effortful process to learn a complex concept, leading to the argument that mapping between the abstract semantic of the number and its verbal tag (i.e. number word) is not automatic. Hence in this account, languages play a central role in learning to count and becoming a cardinal principle knower

Morpho-syntactic properties of number word structure in different languages play an important in learning to count (see § 3.2.1 Transparency of power). For example, compared to English speakers, Chinese speakers can count higher and make fewer errors (Miller et al., 1995; Miller & Stigler, 1987). More recently, Lonnemann and colleagues (2019) compared preschool 5-year-old German and Chinese speakers on a counting and a non-symbolic magnitude number comparison task where the children had to decide on which side of the screen there was the largest set of dots. While the results indicate indeed better counting performances for the Chinese-speaking children than the German-speaking ones, no differences in terms of reaction times were observed on the non-symbolic magnitude comparison task. Schneider et al., (2020), compared 4 to 5 years-old-children speaking more transparent (Cantonese, Slovenian and English) to more opaque languages (Hindi and Gujarati) with several tasks on counting tapping into the successor function, which is an indicator of having acquired the cardinality principle. They found that the more transparent language speakers acquired the successor function more easily than opaque language speakers.

In sum, these studies suggest that learning to count is influenced by language transparency. Nevertheless, the processes itself is sustained by common cognitive mechanisms

[8] Although fingers, body part or an external tool such as small rocks could also be used to store the exact quantity.

across languages. The acquisition of counting and cardinality is facilitated in languages with consistent morphos-syntactic transparency.

3.3.2 Lexico-semantic associations

The quality of association between numbers determines their **semantics**. For example, the numerical distance between numbers (i.e. between 5 and 8 the distance is 3) determines their strength of association. This can be observed behaviourally with the distance effect. The **distance effect** is a cognitive phenomenon that predicts that the distance between numbers affects how they are processed. Initially, the distance effect was observed by slower reaction times in magnitude judgment tasks (i.e. judge which of two numerals is larger) between closer (i.e. 4 vs 5 = distance 1) than distant (i.e. 1 vs 5 = distance 4) pairs (Moyer & Landauer, 1967). Several explanations have been provided for the origin of the distance effect and its origin is still debated today. One of the first explanations for the distance effect was developed on the Triple Code Model and Approximate Number System (ANS) account: since numbers are represented in a mental number line larger representational overlap is expected for closer pairs (Dehaene, 1992; Naccache & Dehaene, 2001). The ANS account is sustained by the finding of a parietal activation, the brain region where the ANS was theorized to be processed (Pinel et al., 2001). For the Discrete Semantic System (DSS) (see § 16 Computational Discrete Semantic System), however, the distance effect is elicited by the linguistic characteristics of number words and their associations, independently from any abstract representations system such as the ANS (Krajcsi et al., 2016; Van Opstal et al., 2008 for another account). Hence, the number distance effect is elicited by the semantic association among number words, like any other words, rather than the association with an external semantic. Nevertheless, both accounts make very similar predictions concerning the distance effect.

Distance effect has also been found in priming studies, leading to a **Priming Distance Effect (PDE)**. In those studies, a numeral precedes the other one sequentially. Hence in PDE, there is a Prime which is followed by a Target stimulus which is relevant for the task. In general, a larger priming effect can be expected if (1) the same stimuli is used as the prime and target, i.e. repetition priming (2) increasing the Stimulus Onset Asynchrony (SOA, i.e. the time between prime and target) weaker semantic priming effect can be expected, and (3) increasing prime duration increases the priming effect (Van den Bussche et al., 2009). The PDE predicts that the closer distance between the sequentially presented pairs, the faster the processing time on the second numeral is (den Heyer & Briand, 1986; Koechlin et al., 1999). Importantly the PDE works also with number words (Reynvoet et al., 2002)

Other lexico-semantic associations that are not discussed here include magnitude, parity and particular numerals such as "encyclopaedic numbers" which can activate associations that go beyond numbers such as 1984 (see Lochy & Schiltz, 2022).

3.4 Chapter Summary

We have reviewed language-dependent influence of number word morpho-syntactic transparencies. How different languages' linguistic accentuation can affect early numerical acquisition (§ 3.1 Language general: Grammar plural markers), how languages' morpho-syntactic transparency differences can affect lexical access (§ 3.2.1 Transparency of power § 3.2.2 Transparency of order) as well as the acquisition of cardinality. We have also reviewed how to experimentally elicit semantic associations of number words (§ 3.3 Number word's semantic).

4 Bilingualism

An intuitive **definition** of bilingualism could be deducted etymologically: composed of the Latin word "*bi*" (two) and "*lingua*" languages, it is hence the ability to use at least two languages. Languages serve the scope of communication between individuals. Note that other means of communication exist, including gestures, emotional tones and inferring the intentional meaning. However, human language differs in that it follows a rule-based organization: the syntax (Friederici, 2017). Bilingualism[9] can be defined as "*[..] the use of at least two languages by an individual or by a community in everyday life situations* "(Grosjean, 2010) or "*[..] the*

[9] Bilingualism is used here as a specific case of multilingualism for more than two languages. I use bilingualism for simplicity by implying the concepts can, to some extent, be generalized to multilingualism. Note also that language is used here in the broader sense including dialects or sign languages. Although some authors argue there are no distinction between languages, rather only idiolects (individual set of words), see (Otheguy et al., 2015). Without entering the socio-political debate of what a language is or not, I use here the pragmatic conventional definition of intelligibility, since the role of language is communication, the exchange of information. If a speaker of a language is severely limited in communicating with the speaker of another language – is unintelligible - then I consider it a different language. German and French, for example, fall under this definition for being different languages.

ability to use two languages for effective communication and learning“ (Bhatia & Ritchie, 2013). Hence the definitions might differ between the *how* and *when* language is acquired. For example, the format in which the languages are learned (i.e. spoken or written), might determine the *how*. And when in life a language is acquired (early or late), might determine the *when.* Both however do not clearly define proficiency, which can vary across bilinguals. The understanding of how bilingualism influences cognitive processes is very **relevant** since it concerns more than half the world's population (Grosjean, 2010). There are different reasons why individuals are bilinguals: from having parents speaking different languages to migrating or moving to a place where another language is spoken. An individual might also become bilingual through school, either when the school language differs from the language(s) spoken at home or when the school proposes a bilingual curriculum. In those cases, the second language (L2) is acquired formally through education (and informally through social interactions) on top of a first language (L1) or home language (HL).

When investigating bilinguals it is important to note their specific cognitive processes which fundamentally differ from monolinguals, i.e. they are not two monolinguals in one (see the opinion paper of Grosjean, 1989). Another important point in investigating bilinguals is that bilingual language profiles can be very heterogeneous and differ among individuals. In this chapter, we hence start by reviewing factors that influence bilingual **heterogeneous** language profiles (§ 4.1 Bilingual heterogeneity). Bilinguals, distinctively from monolinguals, can switch between languages, even within a sentence (i.e. “you should never travel without *une serviette*”). Language switching can also occur when a bilingual is tested in a different language than the one in which the testing content was learned. Then we will see what bilinguals have in common, which **homogeneity** of bilingual cognitive processing is predicted by cognitive models (§ 4.2 Bilingual models). Finally, we will review evidence for and against the proposition that bilingualism elicits some cognitive benefits which are the clear costs identified by research (4.4 Cognitive Cost and Benefits of Bilingualism).

4.1 Bilingual heterogeneity in language profiles

Despite bilinguals being investigated and described as a group – which can conveniently be compared to monolinguals – bilingual language profiles can be very heterogeneous. The heterogeneity is for example represented in such that each language(s) might have been acquired at the same (simultaneous bilingual) or at different times (sequential bilinguals). Each language can differ in the Age of Acquisition (AoA). There might also be different degrees of exposure and use of each language, influencing the relative individual frequency of each language (i.e. frequency increases with more exposure). Bilingual L1 and L2 proficiencies can vary, they can be balanced or unbalanced, high or low. The context of the use of each language also affects bilingual profiles (*i.e.* at home, in school, with friends) or domains (i.e. learning mathematics or writing scientific articles).

Age of Acquisition (**AoA**) is the onset age at which a language has been acquired. AoA can affect language proficiency for a very simple reason that acquiring a language earlier gives more time for use and exposure. However, as we have seen before, interpreters and teachers could even reverse this. This is also the case for migrants when moving to a country where another language spoken might drastically lessen the use and exposure to the first acquired language. Bilingual research oft uses an L2 AoA cut-off to distinguish early from late bilinguals which spans between 3- and 7-year-olds. This kind of cut-off to categorize early and late bilinguals comes from the **critical period hypothesis (CPH).** The CPH states that, due to biological constraints, after a certain age it is not possible to acquire a second language proficiency that is the same as the first language (Lenneberg, 1967). Like late acquisition of language has some dramatic effects already generally for language[10] it might be that similar biological constraints would be underlying equal proficiency in L2 acquisition. Several studies have found decreased proficiency for late L2 learners (Flege et al., 1999; J. S. Johnson & Newport, 1989; Weber-Fox & Neville, 1996), sustaining the CPH and importance of AOA. Hartshorne et al., (2018) collected a very large sample online and found that native-like grammar acquisition in a second language needs immersion learning before the age of 17 years old, after which given the decline in learning ability and time years acquired in learning a new language it is not possible. A strict AoA for CPH such as the existence of a breakpoint in age after which proficiency in L2 can only be lower than in L1 is however criticized. For example,

[10] For example feral children - children who grow up without contact with other humans - are unable to acquire a language proficiently. A famous case occurred in France; a child named Victor was discovered living alone in the woods and taken into custody when he was about 12 years old. Despite 5 years of efforts in teaching him French, Victor was only able to learn to understand “non” and say two words (“lait” and “Oh Dieu!”) (Itard et al., 1891)

different AoA might depend on the language of second language learning and the statistical tools used to calculate the AoA breakpoint (Vanhove, 2013).

Language **proficiency** differs already among monolinguals; hence it can differently affect both bilinguals' languages, such as being high or low. Relative proficiency between both bilingual languages can be balanced (i.e. similar for L1 and L2) or unbalanced. Language proficiency is affected by exposure which affects the relative frequency of each language. Hence proficiency and AoA are hard to disentangle since in most cases the earlier acquired language are also the ones benefitting from most individual frequency of exposure and use. In other words, earlier learned languages are also the ones that bilinguals have used the most in their life, leading to stronger associations from a connectivist perspective. For example, absolute L2 language proficiency can be or become very high for interpreters, foreign language teachers or in the case of second language immersion.

Education can play an important role in language learning. Many individuals become bilinguals through formal education. This is the case for migrants when the home language (HL) differs from the language learned at school and in which individuals can be immerged, in which case bilingualism arises incidentally from the difference between HL and schooling language. However bilingual education can also be planned with **bilingual school curricula.** These curriculum aims, in addition to general education goals such as literacy and numeracy, to teach at least a second language. Bilingual school curricula are however designed and follow several models of bilingual schools such as (Baker & Wright, 2021; Cummins, 2016):

- Content and Language Integrated Learning or Content-Based Instruction (CLIL), where specific contents are taught in different languages (e.g. English math classes in France).
- Two-way Immersion, (also called bidirectional or bilingual immersion) where ideally half minority and half majority speakers are taught together in classes in both languages, hence language is learnt through content.
- European School, International School and Transitional Bilingual Education are mostly private and expensive bilingual school programs. For example in European schools, children can use their native language for learning in primary school but can be instructed in English, French or German.

The evaluation of the effectiveness of these methods and compared with monolingual education is a highly political debate (Baker & Wright, 2021). Both incidental and planned second language education might lead to different profiles. For example, individuals' motivation and emotions (i.e. enjoyment and anxiety) of second language learning (Dewaele et al., 2023) can differently impact second language acquisition in incidental or planned settings. Note that a difference between HL and education language might also result from a different regional minority HL (i.e. dialects) with a national majority language used for education.

Besides bilingual education, other contextual factors specific to bilinguals can affect proficiency. For example language switching cost (**LSC**). LSC occurs when switching between languages, which can occur in the long term, switching the languages between learning and retrieval, as well as in the short term, switching languages between items or sentences. Finally, even the testing linguistic context might play a role in the state of bilingual language profiles as we will see in the Language mode model reviewed below (§ 4.2.4 Language mode). Note that short-term LSC and language mode affect bilingual language status at the moment they are measured.

Finally, bilingual language profiles also depend on the linguistic properties of each language. For example, transparencies can affect second language acquisition. More transparent languages might be easier to acquire and closer linguistic distances between L1 and L2 might also impact language proficiencies (see for example § 3 Language-dependent influences on the exact verbal representation of numbers, for the specific case of numbers).

In sum, bilingualism should be considered rather dimensionally (i.e. early/late, proficient/non-proficient, balanced/unbalanced) rather than categorically (i.e. bilingual or not bilingual). Given the heterogeneity, importantly, not all studies on bilinguals can be generalized to all bilinguals and it is important to carefully control or acknowledge the sample's bilingual language profiles.

4.2 Bilingual models

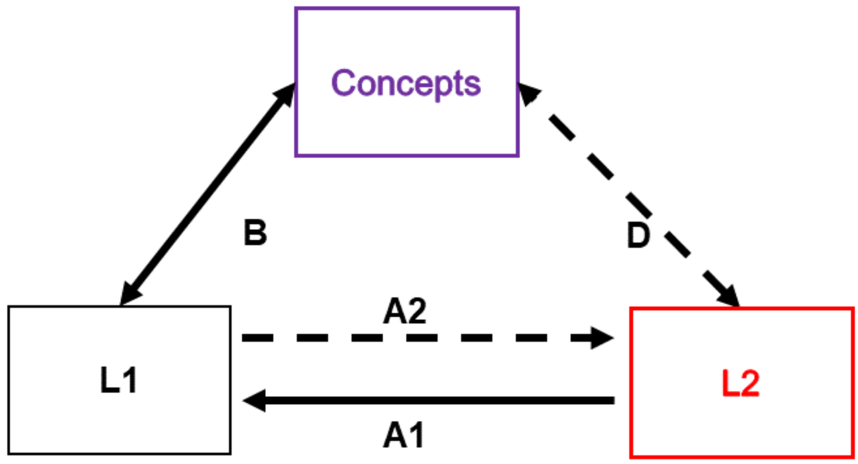
Bilinguals have theoretically double the mental lexicon of monolinguals: one for each language. With at least two words related to similar objects for example and overlapping semantic meanings for more abstract words. The cognitive mechanisms underlying bilingual lexical processing therefore differ from that of monolinguals. For example multiple words with large semantic overlaps it means might be co-activated with the necessity of inhibiting one of the languages. These propositions are derived from the *single network hypothesis* that stipulates bilinguals languages are always (co)activated in parallel (i.e. *parallel activation* see Kroll et al., 2015). For example, the Arabic numeral 5 could co-activate the number representation of “fünf” and “cinq” in the German-French bilingual brain. In the description of the following bilingual models, we will see different explanations for how this competition is resolved. Such that when a German-French bilingual sees a 5 the output /cinq/ leads to remarkably rare involuntary language confusions (i.e. saying /fünf/), (see Gollan et al., 2011 for a study on language intrusions in older bilinguals).

4.2.1 Revised Hierarchical Model (RHM)

The Revised Hierarchical Model (RHM) represented in Figure 7**Figure *5***, proposed that the L2 has weaker semantic connections (i.e. D) but strong connections with the L1 (see A1) (Kroll & Stewart, 1994). However, the connections between L1 and L2 are weaker which suggests an asymmetry that L2 → L1 backward translations are easier than L1 → L2 forward translations (cfr. A1 and A2 in Figure 7. Taking the example of numbers, a German (L1)-French(L2) bilingual would be facilitated in translating “cinq” into “fünf” (L2 → L1) rather than “fünf” in “cinq” (L1 →L2). Translation asymmetries are supported by studies on language switching costs (see § 5.2 Language Switching Cost (LSC)).

Figure 5

Revised Hierarchical Model (RHM)



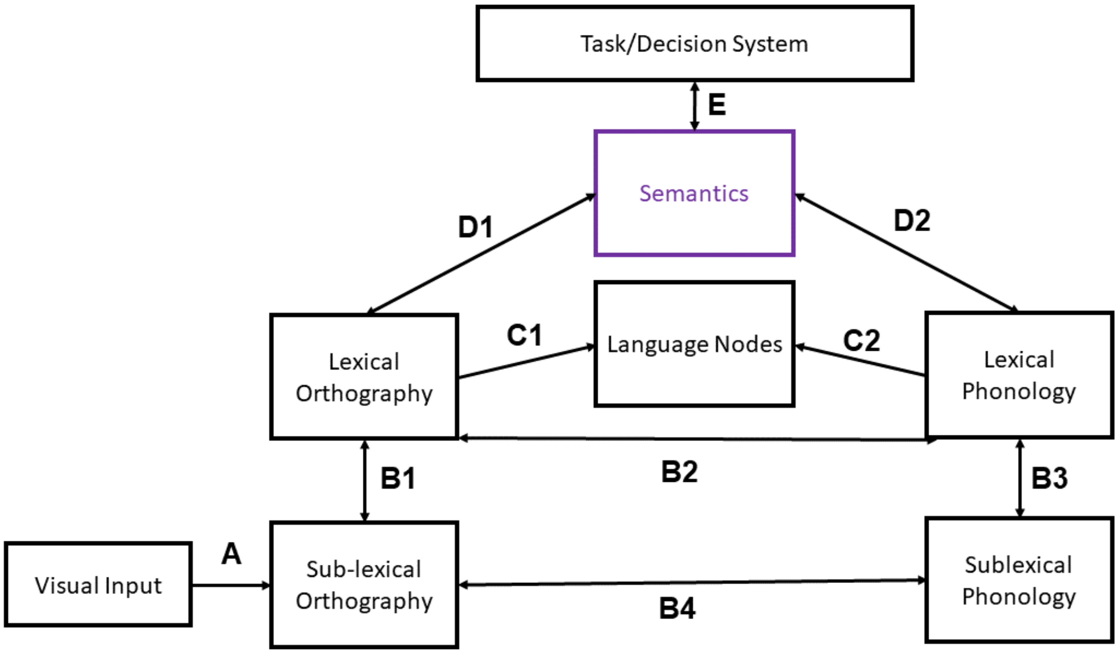
Note: Adapted by the author. A1 and A2 are lexical associations between L1 and L2. L1 to L2 associations are weaker (i.e. A2) than L2 to L1 (i.e. A1). B and D are conceptual (lexico-semantic) associations of both languages. L1 has stronger associations with concepts (i.e. B) than L2 (i.e. D). The RHM is however criticized for postulating separate lexicons for the L1 and L2 (Brysbaert & Duyck, 2010).

4.2.2 BIA+, BIA-d and Multilink

The observation that multiple languages are co-activated in bilinguals has led to the conceptualization of the Bilingual Interactive Activation (**BIA+**) model (Dijkstra & van Heuven, 1998, 2002), which was originally developed for bilingual word recognition. In the model, both languages are integrated into a common lexicon. Hence the BIA+ predicts that lexical activation is language-independent. Words are entered as visual input into a sub-lexical orthography module (see Figure 8, line A). The word is then processed together with a lexical orthography and sub-lexical and lexical phonology module. The two lexical modules are connected unidirectionally (bottom-up) with a language node and bi-directionally with a shared semantic. In addition to this linguistic module whose goal is word identification, there is a decision system that might be affected by the participant’s expectation of the task, see Figure 8.

Figure 6

Bilingual Interactive Activation (BIA+)



.

Note: Adapted by the author. A is the association between visual input and lexical processing network represented by B1 to B4. C1 and C2 represent the association between Orthography and Phonology with the language node. Hence the language is detected at this stage. D1 and D2 indicate lexico-semantic associations with the orthographic and phonological forms. E is then the association of semantics with a task or decision system that can influence the process.

So, for example with the number word "fünf", this is first orthographically and lexically decoded (see B1 to B4), then a language node provides the information that the number word is in German. The word is then associated with its semantic, i.e. meaning (D1 and D2). Finally, depending on the task, the decision system might increase or decrease control over these processes (i.e. E). Hence each word's representation has a resting level of activation that directly depends on the frequency of each language as well as of each word. Hence making the activation level of each word dependent on subjective frequency, recency of use and proficiency (Dijkstra & van Heuven, 2002). Therefore, differences between languages depend on the amount of exposure to each single word, which influences the resting level of activation.

Increasing **exposure** to words in one language compared to the other hence increases their frequency, which in terms of Hebbian learning, means these word representations are stronger. Hence languages with older AoA are in total used less frequently than for young AoA.

Moreover, it explains the difference with monolinguals, since bilinguals have their hours of exposure divided between two languages it consequently diminishes the subjective frequency for both languages (i.e. Gollan et al., 2008). It also explains L2 lexical and semantic costs (i.e. temporal delays) given that the L2 have lower subjective frequencies than the L1 (Dijkstra & van Heuven, 2002).

The BIA + model however is **criticized** for not explaining development-related changes in language proficiency such as second language learning. A developmental model has been suggested by Grainger et al., (2010) to address this model: the BIA-developmental or BIA-d. In the BIA-d, first, there is a lexical association between L1 and L2, then increasing L2 exposure leads to a direct association with common semantics and decreased association with the correspondent L1 lexicon. In other words, the L2 increases its direct and independent association with common semantics with increasing proficiency. Interestingly, the author suggests that for the BIA-d, increased L2 proficiency is realized by increased inhibitory connections of the L1 word orthographic form (reminding the following inhibitory control model, see § 4.2.3). For example, a German (L1) learning French as a second language (L2) will first learn that the number word "cinq" is the same as "fünf". Then, with increasing proficiency in the L2, the number word "fünf" will increase its direct lexico-semantic association with the meaning of this number. Finally, "cinq" will have a strong lexico-semantic association which will be independent from the associations of "fünf". At the same time the association between "cinq" and "fünf" will become weaker.

The BIA+ model has been recently been updated into the **Multilink model** (Dijkstra et al., 2019). This model aims at integrating cognitive and computational models of bilingual word recognition. BIA+ also predicts short-term LSC since the competing language node needs to be inhibited when changing languages between trials (Declerck & Philipp, 2015).

4.2.3 Inhibitory and Adaptative Control (IC and AC)

The inhibitory control (**IC**) model has been developed mainly for language production. Like in BIA+, the model postulates the co-activation of both bilingual languages. The IC model is inspired by action control theories: like actions are actively inhibited by the brain, words in the undesired language are also inhibited (Green, 1998). The IC suggests that different words have different levels of activation, hence inhibition is more efficient for languages with lower levels of activation than higher ones. This is observable in that languages with higher levels of activation benefit from faster processing since the competing language is more efficiently inhibited than the other way around. The difference between the BIA + and IC models is that while in the BIA+ model, the different languages are inhibited on the lexical forms, the IC model inhibits the language tags at the lemma's level. In the IC model therefore there is a Supervisory Attentional System that inhibits and activates lexico-semantic representations. Since a competing word needs to be inhibited in bilinguals, the IC model explains why this group is in general slower than monolinguals in object naming (Mägiste, 1979). Furthermore, an indirect prediction of the IC hypothesis is that bilinguals should be stronger than monolinguals in one of the executive functions: inhibition. This would therefore lead to a cognitive advantage of bilingualism, on cognitive control, which is however a controversial question in the current state of research (see § 4.4 Cognitive Cost and Benefits of Bilingualism).

Adaptative Control (**AC**) is an updated model of the IC model that includes context-dependent factors for bilingual language use (Green & Abutalebi, 2013). Hence, the process described by the IC can adapt to three different contexts: single-language, dual-language, and dense code-switching. The difference between dual-languages and dense code-switching is that during dual-language the language might change between conversations, while in dense code-switching it might even change within a sentence (i.e. "there are *cinq* books in my bag"). In addition, the model suggests there are eight cognitive control processes: (1) Goal maintenance and (2) conflict monitoring as top-down processes necessary for efficient cognitive processing. Then, (3) Salient cue detection (4) Response inhibition, (5) Task disengagement and (6) Task engagement are particularly recruited in dual-language contexts to switch between languages. Finally (8) Opportunistic planning is conceived in the AC model for dense code-switching to plan a sentence mixing different languages and syntaxes. For example, (Abutalebi et al., 2008) investigated German-French bilinguals from a translation school in a picture naming task. They varied the context of the task with either a simple naming context where all pictures of a block were named in a language or a language selection context where a cue indicated in which language to name the picture. They found that the mixed language context activated the left caudate nucleus and anterior cingular cortex which are also involved in non-linguistic tasks involving cognitive control (Abutalebi et al., 2012).

Inhibitory control predicts short-term Language Switching Costs because the language inhibition exerted on the previous trial needs to be overcome when switching to the language of the following trial (Declerck & Philipp, 2015)

4.2.4 Language mode and complementarity

Language context is also an important factor to consider when evaluating bilinguals. Grosjean (2001) postulates the existence of a language mode: a bilingual's ability to tune to different languages depending on the context. Once one language mode is activated this language is easier to retrieve. For example, a German-French bilingual in a class in French would be in a French mode which will facilitate the lexical retrieval and comprehension of French relative to being in a German mode. Hence in the language mode, the level of activation of each language depends on the perceived and precedent language context. Complementarity is another notion that was added later to the theory (Grosjean, 2010). This notion states that languages might also be specific to certain domains (i.e. social, formal, etc.) and that the most proficient language is the one that covers most domains. Hence this also suggests that a bilingual's language proficiency might be domain-specific.

4.2.5 Summary of the models

All the reviewed models[11] imply the associations of bilingual languages can differ. This is for example postulated in the RHM with different weights for each language. For BIA+ language differences emerge by difference of exposure impacting the level of activation and association of or specific to words. Hence each word's relative frequency predicts the level of activation specifically from the language. The IC model focuses on language selection, suggesting an active inhibitory mechanism to prevent unwanted interference for co-activated languages. Both the BIA + and IC/AC models underline the importance of the task of how bilinguals process languages. Finally the "language mode" puts forward the importance of context in bilingual language processes.

4.3 Bilingual brains

The brain is responsible for integrating and managing the cognitive mechanisms underlying bilingual language processing which has been modelled in the previous chapter.

[11] Note the existence of other models such as the Bilingual Model of Lexical Access (BIMOLA) (Grosjean, 2008), the Self-Organizing Model of Bilingual Processing (SOMBIP) (Li & Farkas, 2002), and the Bilingual Simple Recurrent Network (BSRN) model (French & Jacquet, 2004) which are not described here..

Support and inspiration for these models come from neuropsychological and neuroimaging studies.

Insights into how bilingual languages are represented in the brain come from **direct cortical stimulations or neuropsychological case studies**. Fernández-Coello et al., (2016) investigated 13 bilingual patients with direct cortical electrical stimulation mapping on awake patients while undergoing neurosurgical procedures. The spoken languages were diverse but could be classified into early acquired (both languages acquired after age 7) and later acquired. During a picture naming task, they found many overlapping brain areas when naming in both languages, but more areas were involved for patients with early than late L2 acquisition. Fabbro et al., (2000) describe a bilingual Friulan-Italian (L2) patient who after a lesion of the left frontal lobe and right anterior cingulate was pathologically switching both languages between sentences, despite being instructed and aware to speak only in one language. Hence this confirms both overlapping representations of languages which are in turn modulated by AoA.

Evidence for **brain overlaps in both languages** on both structural (i.e. grey matter) and connectivity (i.e. white matter) is confirmed by evidence from meta-analyses of fMRI neuroimaging studies (Abutalebi et al., 2001). Hence supporting the IC and BIA+ models' postulate of bilingual co-activations of both languages.

The brain overlaps between L1 and L2 seem however increase for L2's early AoA, high exposure and high proficiencies (Indefrey, 2006) as shown by several **meta-analyses** on the topic. For example, lower L2 proficiency involves a greater activation in pre-frontal areas and the caudate nucleus (Abutalebi, 2008; Li et al., 2014). In a meta-analysis, Liu & Cao, (2016), found that late bilinguals recruited more brain regions for processing the L2 than early bilinguals. Hence, depending on AoA, L2 processing recruits more brain area than L1, suggesting here more cognitive demand and more effort during L2 processing. On the contrary, early bilinguals displayed more activation of the left fusiform gyrus than late bilinguals during L1 processing, a brain part that is also associated with orthographic processing. with regards specifically to the lexico-semantic processing brain network, bilinguals' L1 overlaps that of monolinguals (Sulpizio et al., 2020). This meta-analysis further found larger activation for late than early L2 learners of the caudate nucleus, which the authors interpret as a monitoring process for L2 semantic information. In sum, brain overlaps in L1 and L2 processing seem to depend on AOA.

Longitudinal evidence of structural modification after L2 acquisition has been found in brain imaging studies **within subjects** after having learned a new language. Stein et al., (2012) investigated L1 English who learned German (L2) in Switzerland. After 5 months of learning and exposure to German, they found increased grey matter volume in the left inferior frontal gyrus and left anterior temporal lobe. Mårtensson et al., (2012) investigated L2 learners in the context of intense language courses for interpreters. They found increased grey matter volume in the left middle frontal gyrus and left superior temporal gyrus. The increase in grey matter in the left inferior frontal gyrus found in the previous study was found only when comparing the interpreters with a control group and not longitudinally. Hosoda et al., (2013) found that after 3 months of training and exposure to English, Japanese speakers underwent a reorganization but exclusively and in the right hemisphere. This reorganization consisted of an increase in the connectivity between the right inferior frontal and superior temporal cortex's connectivity and an increase in the volume of the right inferior frontal cortex. In sum, the three longitudinal brain imaging studies reviewed here suggest a general involvement of frontal and temporal structures in second language acquisition. They also underline the variability of results between studies both in terms of structures (i.e. inferior frontal gyrus) as well as lateralization (i.e. right vs left).

Studies have been carried out on structural brain changes **between bilinguals and monolinguals**. Functional Magnetic Resonance Imaging (fMRI) research has found a larger grey matter in the left inferior parietal module for early bilinguals than for late bilinguals and monolinguals (Mechelli et al., 2004). Bilingualism, as reviewed by (Li et al., 2014) also makes the white fibres of certain networks remain more integer with increasing age. These findings have been put forward to sustain the "cognitive reserve" hypothesis (Gold et al., 2013; Luk et al., 2011). Hence increasing L2 proficiency and earlier AoA is reflected in structural changes in the brain.

In sum, bilinguals display both structural and functional changes in the brain. These changes are found in different frontal and temporal areas. In general, late and less proficient languages are found to activate more brain areas than early and more proficient languages leading to less overlap for processing both and suggesting the late acquired languages are processed less efficiently than early ones.

4.4 Cognitive Cost and Benefits of Bilingualism

In the previous chapter, we have seen that bilingual brains differ in terms of structure and connectivity from monolinguals. Those structural differences are likely due to differences in how bilinguals process information compared to monolinguals, which is - at least theoretically - expected (see § 4.2 Bilingual models). Note however that the cognitive costs and benefits might vary across individual factors such as the bilingual language profile, societal factors such as how specific language knowledge is perceived, and finally linguistic commonalities or differences between the languages. The question of cognitive cost and benefits in bilinguals is also important in numerical cognition since several cognitive functions are related to those. Executive functions for example correlate with numerical and mathematical thinking (Bull & Lee, 2014; Clement et al., 2016; Coolen et al., 2021). Some studies have indeed found a bilingual advantage for bilinguals mathematical performances compared to monolinguals (Hartanto et al., 2018; Kempert et al., 2011; Marian et al., 2013; Prior et al., 2015a; Stocco & Prat, 2014), suggesting that the cognitive advantage in executive functions might facilitate mathematical reasoning in bilinguals.

4.4.1 Cognitive Benefits

The possibility of a **bilingual cognitive advantage** is a highly debated and controversial topic (Lehtonen et al., 2018). The hypothesis was originally generated from the reasoning that bilinguals train their executive functions when inhibiting competing languages or switching between languages, hence resulting in a domain-general cognitive advantage over monolinguals. Furthermore constituting cognitive reserve protects against cognitive decline with ageing (Bialystok et al., 2004). This account has been sustained by several meta-analyses: one finding positive results for bilinguals across several cognitive domains such as meta-cognition, and meta-linguistic (Adesope et al., 2010) and another one specifically for verbal working memory (Grundy & Timmer, 2017).

However, since the first meta-analysis **did not find any advantages** for bilinguals (Hilchey & Klein, 2011), several authors have criticized the initial finding of a bilingual advantage to be explained by a publication bias[12]. Gunnerud et al., (2020) and several meta-analyses have not found any bilingual advantages. For example, de Bruin et al., (2015) found that conference abstracts with positive results had more chances to be later published. Also, when meta-analyses controlled for publication bias the cognitive advantage effects disappeared both when looking at adult (Lehtonen et al., 2018) and children populations (Lowe et al., 2021). Lehtonen et al., (2018), further observed that studies with large positive effect sizes were those with smaller samples. Secondly, they pointed out that if the bilingual advantage is domain-general it should be found in multiple components of the executive functions since they correlate (i.e. Miyake et al., 2000). In their meta-analyses, they investigated different sub-components of executive functions without finding any advantages. Hence confirming initial negative results about the bilingual advantage using several tasks tapping on executive functions (Paap & Greenberg, 2013). Large-scale studies (i.e. more than 4500 and 11000 participants) also failed to find a bilingual advantage both in children (Dick et al., 2019, see also Goldsmith et al., 2023) and adults (Nichols et al., 2020). Finally, in a meta-analysis of several of the meta-analyses mentioned above (i.e. a meta-meta-analysis) Paap et al., (2024) found that studies from a single lab were the main moderator for the different results. Hence, they concluded that bilingualism, like other cognitive training, does not result in far transfer, *i.e.* a general bilingual advantage.

A cognitive advantage could originate from some bilingual's **habit of switching between languages**, rather than from being bilingual *per se*. In a study by Verreyt et al., (2016), highly proficient Dutch-French bilinguals were recruited in Bruxelles, Belgium. Two subgroups were made based on whether participants were used to switching very often between languages or not. They then did a flanker and Simon task, two well-established tasks which need to inhibit salient interfering information. The results showed a larger cognitive advantage for the bilinguals who switched very often compared to those who did not. In a further longitudinal study, Woumans et al., (2016) did a longitudinal comparison between 5-year-old children going to monolingual or bilingual school. In the first measures, the children did not differ on verbal fluency nor on socio-economic status, however, 1 year later the bilinguals

[12] A publication bias occurs when only positive results of a bilingual advantages are published, while results not finding any advantages are not published and therefore not visible.

scored higher on intelligence tests than monolinguals. Hence a bilingual cognitive advantage could be explained by a bilingual's switching experience rather than being bilingual *per se.*

4.4.2 Cognitive Costs

Having an integrated vocabulary including two languages can lead to two costs compared to having only one language. The first cost regards **vocabulary** size: compared to monolinguals bilingual language receptive vocabulary is smaller (Bialystok et al., 2010). Smaller vocabulary in bilinguals is relative, given that it concerns only one of the bilingual languages, when both bilinguals' languages are compared the vocabulary is larger than monolinguals (Bialystok & Feng, 2009). Bilinguals also have more interference in decision tasks and tip-of-the-tongue phenomena (Bialystok, 2009). This might be due to delayed vocabulary acquisition in bilinguals compared to monolinguals, such as in a large-scale study on Spanish-English bilingual children in Miami (Oller et al., 2007). Lower vocabulary is found in adult bilinguals with immigration backgrounds in the US, despite the average age of arrival in the U.S. being before puberty 10 years old (Portocarrero et al., 2007). In contrast, another study comparing Canadian English monolingual to Indian Tamil-English (L2 acquired at the age of 6) bilinguals living in India, found equal vocabulary sizes in both groups (Bialystok et al., 2004).

Slower bilingual L1 responses than monolinguals have led to the **bilingual lexical deficit hypothesis:** bilinguals in their L1 should be slower than monolinguals. An explanation for the lexical deficit is that it is cognitively more effortful to find words in a larger than a smaller repertoire, which is supported by several studies. For example, in picture naming and word reading tasks, bilinguals in their L1 showed larger activation than monolinguals in several parts of the left hemisphere: dorsal precentral gyrus, pars triangularis, pars opercularis, superior temporal gyrus and planum temporale (Parker Jones et al., 2012). Those increased activations in bilinguals were interpreted in that bilingual word retrieval is more effortful than for bilinguals. Another study investigating picture naming in three groups of Spanish monolinguals and highly proficient Spanish Spanish-Catalan and Catalan-Spanish bilinguals found that the bilingual groups were slower than the monolingual group. Importantly the gap indicating faster monolinguals than bilinguals' L1 remained across repetitions (Ivanova & Costa, 2008). (Ivanova & Costa, 2008) compared Spanish-Catalan (L2 AoA 5 years old), Catalan-Spanish (L2 AoA 5 years old) bilinguals studying in Barcelona and Spanish bilinguals studying in Madrid in a picture naming task in Spanish. Both bilingual groups were found to be slower in lexical retrieval, hence indicating a cost for lexical retrieval for bilinguals both in L1 and L2. Similarly, De Bruin et al., (2016) investigated an older (i.e. > 60 years old) Scottish sample who learned Gaelic as children but either actively used Gaelic with English or used only English (inactive bilinguals), in comparison to a third group of English monolinguals. These three groups underwent a word-picture matching task where they had to press if the word under a picture matched or not the task. The results indicate that the Gaelic bilinguals were overall slower than monolinguals, with inactive bilinguals having reaction times in between the faster monolinguals and slower active bilinguals. Hence suggesting that bilingualism slows down lexical access and sustaining the bilingual lexical access deficit.

The IC, AC and BIA+ models have in common that **L1 and L2 vocabularies are integrated into a common network.** So how does the brain accurately select between two semantically related lexical items (i.e. words) from different languages? For the weaker link hypothesis (Michael & Gollan, 2005), language selection is resolved by frequency: languages that are retrieved more frequently (L1) are easier to access than words that are retrieved less frequently (L2). However, the individual L1 frequency is in general lower than for monolinguals since monolinguals only practice one language while bilinguals practice two languages. For the IC and AC models, language selection is operated by inhibiting the undesired language: since the L1 has a stronger association, it is more effortful to inhibit than the L2, leading to slower lexical retrieval of the L2. However, the L1 too is slower to retrieve compared to monolinguals, since the competing L2 vocabulary needs to be inhibited. This corresponds to the prediction made by the RHM: L1 is faster to retrieve than L2. Hence all three models predict slower L2 than L1 but also slower L1 than monolinguals.

5 Bilingual numbers

A substantial difference between monolingual and bilingual number processing resides in that bilinguals have two different languages to process numbers. This **bilingual specificity** fundamentally affects bilingual number processing in comparison to monolinguals. This distinction has already been underscored in the context of general language processing (Grosjean, 1989). The representation of numbers in different languages, both in terms of lexical representations and their associations to visual and semantic representations, presents a challenge in terms of complexity when attempting a description of these phenomena. This complexity can be decomposed into different factors. One of those factors is having two verbal representations instead of one, each of which might linguistically differ such as in morpho-syntactic structures, as we have seen for monolinguals (§ 3 Language-dependent influences on the exact verbal representation of numbers). Another important factor, compared to monolinguals is that bilinguals have acquired two languages, leading to (even more) heterogeneous language profiles than monolinguals (see § 4.1 Bilingual heterogeneity). A final factor is that bilinguals can also switch between their languages (which is impossible for a monolingual). Concerning language switches we need to define a second distinction: long-term and short-term switches. Long-term switches can for example occur between learning and testing, such as when training in one language and testing in another one. Short-term switches occur between trials (i.e. within a test), that is when after having given the response in one language the following response is requested in another language. Note that it can also occur situationally or contextually, such as when a test is done in a different language than the one that has been spoken or activated before doing the test. Both short- and long-term switches lead to language switching costs (LSC): performances get worse after switching than remaining in the same language. Finally, on top of the complexity of each of these single factors, the overall phenomenon's complexity can increase due to interactions among each factor. Given the higher prevalence of bilinguals than monolinguals, as highlighted by Grosjean, (2010) along with the significant role of mathematical education in forecasting school success, future adult's socioeconomic level (Duncan et al., 2007; Ritchie & Bates, 2013), it is important to bring light into our understanding despite the potential complexity of these factors and their interactions on the representation of numbers.

In the following, we will start by reviewing the effect observed on bilingual learning and solving mathematics (§ 5.1 Bilingual). Theoretical cognitive models can be a way to simplify our view and increase how bilinguals represent numbers (§ 5.3. Model of bilingual number representations). We will then see empirical studies investigating and comparing different aspects and associations of these representations: lexical (§ 5.4 Bilingual lexical representations of numbers) and semantic (§ 5.6. Bilingual lexico-semantic association of numbers).

5.1 Bilingual mathematics

The **formal acquisition of mathematics** occurs in schools, for some students, it can occur in a different language than the one spoken at home. For others, the language changes due to a change in curriculum such as moving across language borders. Finally, a language change can also occur within the same school curriculum as in Luxembourg where the language switches from German to French as part of an education policy fostering the learning of both languages. All these examples involve the switch to a new language which needs to be first acquired (or at least consolidated) for the student to access the content and requirements of scolarization. Here we will focus on the effect of different types of changes on the learning of mathematics, such as for word problems. We will use the following terminology to design the language status: L1 is the self-reported most dominant first language hence the most proficient and in most cases the first acquired language (in particular for studies with children). LM is the language of learning mathematics, this depends on the school curricula, while for most students in a monolingual education setting it remains the same, it can also change, such that there is an LM1 and LM2. As we will see some authors also use the term LM+ to design the language in which mathematics are learned or trained, in contrast to the LM-. Finally, there is also the home language (HL), the language spoken at home, this terminology is mostly used in the context of migration, where the children are schooled in a different language than the language spoken at home. These language distinctions are very important and relevant for students since bilingual students have a higher risk of being retained compared to their monolingual classmates (Baumert & Schümer, 2002). Also because teachers tend to believe that the teaching of mathematics is independent of general language (Fernandes, 2023), hence underestimating the impact bilingualism might have on the learning of mathematics.

Hereafter we will see how bilingual language skills are related to advanced mathematical and arithmetic abilities.

5.1.1 Bilingual advanced mathematical skills

General language abilities in the language of scolarization are fundamental for learning, even for mathematical skills. For example, a link between language and mathematics is found for more advanced mathematical skills such as geometry and fractions for both L1 and L2 (Kleemans & Segers, 2020; Vukovic & Lesaux, 2013). **Word problems** involve a verbal narrative representing real word situations, requiring a mathematical equation to be solved (Greer et al., 2002). For example "Thom has 5 euros in his pocket, he spends 3 to buy bread, how many euros remain in his pocket?". To be solved, a word problem requires the understanding of the context, and select the relevant information to create a mental model of the problem to be solved (Verschaffel et al., 2000). Since word problems are mostly presented in written form, students' language comprehension predicts the success of their solving (Fuchs et al., 2006, 2015). Hence the proficiency attained in biliterism, being able to read two different languages, and bilingualism might differently predict the success in word problem solving compared to monolinguals and in each language. A study compared Turkish-German bilinguals with German monolingual children going to German-speaking schools. The results show that the Turkish-German bilinguals were better at solving the problems in German than in Turkish. This suggests they were advantaged in solving the problems in the same language as they learned mathematics (LM German) than when they had to switch to their home language (HL) Turkish which can be explained by LSC. For the German word problems, the bilinguals however had lower performances than German monolinguals. Hence this suggests that language mastery (fostered for monolinguals by having the HL = LM) is important for solving word problems. However, for more complex problems that were presented with distractors, highly proficient bilinguals had a relatively better performance than monolingual peers (Kempert et al., 2011). Hence suggesting that bilinguals might nevertheless have a relative advantage regarding attentional control which would in turn enhance their performance in math problems which are particularly demanding in this regard.

Other studies compared bilinguals solving word problems in their L1 to their L2, with **contrasting findings**. Word problems are harder to solve for students whose problem is not written in their HL (Greisen et al., 2021). While some studies find better performances in their L1 (i.e. HL) compared to their L2 (Bernardo, 2002; Bernardo & Calleja, 2005 on Philippino-English bilinguals), other studies showed equal performances in both languages (Secada, 1991 on Spanish-English bilinguals). When the LM is the L2, i.e. a new language that is learned at school, it can be expected that the role of LM (=L2) increases with increasing grades, since exposure and consolidation of the L2 increases with each school year. Hence, while first-grade bilinguals are less effective in word problems presented in their L2, second-grade students see their effectiveness improve (Ester et al., 2021). The relationship between L2 word problems and working memory performances in Spanish (L1)-English(L2) bilinguals are mediated by L1 word problem solving and vocabulary level (Swanson et al., 2022). This relation with vocabulary seems to be specific for bilinguals, as suggested in a study showing that third-grade English learner bilinguals had a lower level in mathematic vocabulary than monolingual peers (Powell et al., 2020).

In sum, these studies show that bilingual's main difficulty with word problems is mediated by knowledge or proficiency in the language in which they are presented. Their executive and attentional functions are not affected and might even be better for bilinguals than monolinguals. In practice, they indicate that language mastery along with language consolidation is necessary for solving word problems. This poses a challenge for multilingual contexts such as in Luxembourg or linguistically diverse classes.

5.1.2 Acalculia and dyscalculia

Acalculia and dyscalculia are two disorders that specifically affect respectively the processing and learning of numbers, mathematics and number sense. While acalculia is a neuropsychological disorder which appears after a brain lesion, acalculia is a neurodevelopmental disorder (American Psychiatric Association, 2013; World Health Organization, 2019)

Following brain lesions, some patients have been found to have specific disorders concerning number processing, and **acalculia**. A German-Greek bilingual has been described to systematically making inversion errors when writing in German, while the inversion errors were absent in Greek (Proios, 2002). For the diagnosis of **dyscalculia**, the two main diagnostic manuals for mental disorders DSM-5 (American Psychiatric Association, 2013) and ICD-11 (World Health Organization, 2019) have as criteria that the learning difficulties should not be explained by a lack of language of math (LM) instructions understanding. Hence underlying the importance of considering the language profiles for diagnosing dyscalculia, such as the LM and HL. Interestingly dyscalculia also affects language skills already in monolinguals (Chow et al., 2021; Forsyth & Powell, 2017; Powell et al., 2020). Increasing general bilingual proficiency in both languages has been shown to longitudinally increase math and working memory performances in Spanish(L1)-English(L2) bilinguals (Swanson et al., 2018). Comparing monolingual and bilingual English learning students with math difficulties shows that bilinguals underperform monolinguals on mathematical vocabulary, while no differences were found for word problems (Powell et al., 2020, 2022). In sum, bilinguals might be at risk of being **misdiagnosed** with dyscalculia due to lower language proficiency (Ugen et al., 2021).

5.1.3 Bilingual arithmetic

Despite that arithmetic skills might be held as independent from the language in popular beliefs, they are related. Similar results for word problems have been replicated when bilinguals solve simple and complex arithmetic. In other words, the memory traces for number facts are stored in a **language-specific way** (Dehaene, 1992; Frenck-Mestre & Vaid, 1993). Arithmetic problems can be solved by rote retrieval from long-term memory of the results or by other strategies such as repeated addition, in particular early in development. As we have seen above in § 2 Models of numerical representations, there are theoretical divergences about the question of how arithmetic facts are represented. For the Triple Code Model, for example, mathematical facts are stored in the language in which they are learned (Dehaene & Cohen, 1995, see also § 2.2 Triple Code Model (TCM)). In the Encoding Complex Model (ECM) bilinguals have different representations of arithmetic facts in each language, hence arithmetic facts are in different verbal stores for each language (or formats) (Campbell & Xue, 2001, see also § 2.3 Encoding Complex Model (ECM)). This has for example been experimentally demonstrated in that participants had worse performances solving arithmetic with articulatory suppression (i.e. repeating a syllable aloud while doing the task) (Moeller, Klein, et al., 2011). The resolution of arithmetic problems such as “2 + 9” or “6 x 8” can involve different strategies that change through development. For example, very young children might rely on counting (i.e. 9 + 1 + 1), and older might use algorithmic computation such as decomposing the problem (i.e. (6x4) x 2) to direct memory retrieval (i.e. the solution is known by heart) (Ashcraft, 1982). All these strategies rely on verbal components such as verbal working memory and lexical retrieval of arithmetic facts for long-term memory, they might therefore not automatically transfer across languages for bilinguals.

A large proportion of bilingual individuals indicate a preference for solving mathematical problems in their **first language of acquisition (L1/HL)** (Kolers, 1968) as revealed in questionnaires (Dewaele, 2007) and in Luxembourg too (Martini, 2021). This preference is reflected in cognitive costs for solving arithmetical problems in the self-reported “non-preferred” language for vocal answers to arithmetic presented in digits (Marsh & Maki, 1976), presented auditory (McClain & Huang, 1982) or presented in number words (Frenck-Mestre & Vaid, 1993). However, the subjective nature of “language preference” can result in circular reasoning to explain this phenomenon, i.e. better performances are found for the preferred language because it is the preferred language. Hence later research has used more objective criteria for determining a bilingual’s language status such as the most dominant first language (L1), language of learning mathematics (LM) or the home language (HL), as we have already seen above.

Worse performances for solving arithmetic in L2 have been observed in **large-scale assessments** and classrooms compared to monolingual students solving them in their L1. Lower performances when the HL differs from the LM have been confirmed by several large-scale studies (Beal et al., 2010; Greisen et al., 2021; Heppt et al., 2015; Martin et al., 2014; Ugen et al., 2013). These difficulties in mathematics for children learning mathematics in an L2 (*i.e.*

when the HL is not the LM) are in part explained by lower instructional language comprehension (Kleemans et al., 2018; Kleemans & Segers, 2020; Peng et al., 2020). For example, specific language skills in mathematics are linked with mathematical skills (Forsyth & Powell, 2017; Purpura & Reid, 2016) and numerical knowledge (Purpura et al., 2017). Greisen et al., (2021) investigated data from a large cohort of 3rd grades (i.e. > 10'000) undergoing the national school monitoring test of Luxembourg (i.e. EpStan, Epreuve Standardisé). They found that math performances were predicted by student's reading comprehension in German, the language of mathematical instruction (LM). Importantly, the effect was stronger for children with a different HL (i.e. Portuguese) than German or Luxembourgish. One study however has not found a difference between bilinguals and monolinguals, after controlling for language background (Hartanto et al., 2018). C. Xu, Di Lonardo Burr, et al., (2022) investigated Students in Canadian with the majority having English as HL who either were in French immersion programs or English programs. Word problem and transcoding tasks that rely on language were related to English proficiency, but tasks they described as having less language demand such as arithmetic and number line estimation did not. Compared to their L1 peers, L2 10 to 12-year-old Dutch speakers have lower scores for higher mathematic problems such as geometry and fractions (Kleemans et al., 2014). This is reflected in lower verbal reasoning in general and poorer language skills (Kleemans & Segers, 2020; Mancilla-Martinez & Lesaux, 2010).

In sum, these studies show a strong relationship between language and mathematical abilities. However, while these studies identify the impact of language dominance on doing math, these designs confound general language and the language of learning mathematics (LM) (Spelke & Tsivkin, 2001a). For instance, the relationship between math skills and L2 language could be confounded by other factors such as language proficiency (Hoff et al., 2012), reading proficiency (Heppt et al., 2015) and socio-economic status (B. W. Sarnecka, 2017).

Nevertheless, there is also a direct relationship between the **bilingual's language status** in which arithmetic is solved and performance. These studies are hence mostly within-subject comparing bilingual performances in both languages. Arithmetic problems are solved faster and with fewer errors in the L1/HL than in the L2. In the US, monolingual and bilingual children underwent a chronometered mathematical fluency task with Arabic numerals (i.e. they had to solve as many problems as possible in a given time) and verbal fluency in English and their HL. The results show that while monolinguals' math fluency correlated with English fluency, bilinguals' math fluency rather correlated with the language fluency measure in their HL (Atagi & Sandhofer, 2023).

Prior et al., (2015) investigated Arabic-Hebrew(L2) and Hebrew-English(L2) bilinguals with different arithmetic problems which were either presented visually (i.e. in Arabic numerals) or auditorily in the L1 or L2 (acquired on average at around 7 to 9 years old). Arabic and Hebrew are written right to left, but only Arabic number words are written in ten-unit form, hence number words are only inverted in Arabic. Their results show that while Hebrew monolinguals preferred problems presented in the non-inverted form (matching their language), the Arabic-Hebrew bilinguals’ performances did not differ in the order of the language of presentation. Hence suggesting bilinguals can flexibly process arithmetic problems whether inverted or not. English-Spanish balanced simultaneous bilinguals did an arithmetic verification task while being recorded with EEG. The ERPs resulted in an N400 component (i.e. a negative amplitude around 400 ms after stimulus onset) in both languages. The N400 is related to semantic processing and was therefore interpreted as a parallel semantic access to the arithmetic problems between both languages (Cerda et al., 2019). In a similar vein, Spanish-English bilingual teachers either learned mathematics in English or Spanish and either taught mathematics in English or Spanish. They had an earlier N400 peak in their teaching language, independently from the language in which they had learned mathematics (Martinez-Lincoln et al., 2015). Hence this study shows that initial networks optimized to retrieve arithmetic facts in one language (i.e. the language of learning mathematics) can – given enough experience (i.e. 9 years in the experiment) – flexibly optimize semantic processing in another language. Hence these two ERP studies suggest that lexico-semantic associations are plastic and can change with intense and ideally early practice in another language.

5.2 Language Switching Cost (LSC)

Differently than monolinguals, bilinguals can also switch between languages. These switches can occur both in the long term such as between learning and retrieval and in the short term such as changing language between the items within a test or between the trials of an experiment. Multiple experimental evidence has shown that switching between languages, both in the long- and short-term involves a Language Switching Cost (LSC). Hence LSC is a cognitive cost explained by changing (i.e. switching) languages compared to when remaining in the same language. LSC, particularly long-term as we will see, could be a part of the explanation for the difficulties in reasoning and doing arithmetic in an L2 compared to the L1 described in the previous chapter. Note that LSC are not specific to numerical cognition, they are found in general language (Marian & Fausey, 2006) and engage many brain areas mainly located in the left temporal lobe (see Luk et al., 2012 for a meta-analysis).

Long-term LSC have been found when the language switches between training and testing, and short-term LSC are found within tests or experiments when a language is switched between items or trials. An example of long-term LSC has been demonstrated by Spelke and Tsivkin (2001b), who trained Russian (L1)-English(L2) bilinguals in solving arithmetic in one language or the other. The participants were faster in the trained language independently if it was the most dominant (L1) or not (i.e. L2). More recently, several studies have shown that participants have a cost when switching language in the long term between learning and testing arithmetic (Grabner et al., 2012a; Hahn et al., 2017, 2019; Kempert et al., 2011; Saalbach et al., 2013; Volmer et al., 2018). The effects of LSC are observable in concrete classroom examples such as for Philippino home speakers who learn mathematics (LM+) in English. Those students are better at solving arithmetic in English (LM) than in Philippino (HL/L1), despite this being their most dominant language (Bernardo, 2001). The gain for solving arithmetic in the LM+, the language in which mathematics is learned, seems to be even stronger than the gain after being trained in a certain language. Austrian adults who attended their school in German were trained to solve arithmetic problems either in English or German. The test consisted of an arithmetic verification task with arithmetic problems presented in number words (i.e. "two times four = eight"). The results indicate that the gain when tested was stronger in German, independently of the language they were trained in (Kraut & Pixner, 2022). In the brain, long-term LSC in arithmetic has been found for English-Chinese bilinguals, who activate different brain areas depending on the language (and script) in which those arithmetics facts were trained (Venkatraman et al., 2006).

Short-term LSC, have been found when switching languages within the trials of an experiment. In general, when a trial is preceded by a trial in the same language it is solved faster than when a trial is preceded by a trial in another language. For example, when a German-French bilingual names the Arabic numeral 5 "cinq" in a first trial and then 4, it will be on average slower to name it "vier" than "quatre" (Jackson et al., 2001, see Declerck and Philipp, 2015 for a review). Short-term LSC are well documented in the general language literature and are also found for two-digit numerals (Contreras-Saavedra et al., 2020, 2021).

In sum, the worst performances observed when the LM is not the HL in bilinguals are in part explained by language mastery (i.e. reading and speech) as well as lower socioeconomic status since it mostly occurs in migrant children. Part of those worse performances can however also be explained in that mathematical knowledge in one language (i.e. HL) does not automatically transfer into another language (i.e. LM), engendering an LSC.

LSC and L(M)2 costs are also found with **brain imaging techniques** such as electroencephalogram (EEG) and fMRI, giving further insights into the neurocognitive mechanisms at their origins. Salillas & Wicha, (2012) investigated fluent Spanish-English bilinguals who learned mathematics in one of both languages (LM+) with EEG. In this study, simple multiplication problems were presented auditorily in both languages and participants had to judge if they were true or false. Participants were faster in judging problems in the LM+ and the averaged EEG signal (i.e. Event Related Potential, ERP) showed a larger N400 response (i.e. negative microvolt peak 400 ms after stimulus onset). Since the N400 component is oft associated with semantic processing, the authors concluded that it reflected better semantic processing of the problems in the LM+. Importantly, stronger N400 in the LM+ were found independently of the language in which the problems were presented (i.e. English or Spanish) and of language dominance (i.e. L1 or L2).

A two-digit number comparison task undergone by Spanish-Basque bilinguals (Salillas & Carreiras, 2014) found that the LM modulated the **number distance effect**. The distance effect arises from the semantic associations of different numbers (i.e. closer numbers such as 4 and 5 have stronger associations than more distant ones such as 1 and 5), hence again suggesting the LM enhances the semantic processing of numbers. On the contrary, when investigating early simultaneous children bilinguals (i.e. who acquired both languages before 3 years old) Cerda and colleagues found no difference in the EEG signal related to accessing arithmetic facts in both languages (Cerda et al., 2019). Also, similar ERP responses in the LM+ and LM- are found in teachers teaching in their LM- (Martinez-Lincoln et al., 2015). Hence suggesting that this disparity in processing arithmetic in the language they have been learned or in another language can be levelled by either intensive training or early acquisition of both languages.

fMRI studies have further confirmed the effect of language dominance on arithmetic in the brain, showing **larger brain activations** for the L2 than L1 (Lin et al., 2012; Wang et al., 2007). Specifically, the temporal regions are more recruited in the LM1, suggesting a direct semantic retrieval compared to when the arithmetic problems are solved in the LM2. The LM2 however recruited brain networks related to more generic cognitive resources (Van Rinsveld et al., 2017). These results are in line with fMRI studies on general language showing larger brain activations for the L2 than L1 (Liu & Cao, 2016). Larger brain area engagement for the resolution of arithmetical problems suggests less efficient neurocognitive processing in the L2 than in the L1.

Therefore this experiment suggests that arithmetic facts are language-dependent, hence changing the language between learning and testing involves a cognitive cost (LSC). They also underline the importance of the learning language of mathematics (LM) on arithmetic performances.

Finally, another language factor that can affect how arithmetic is processed is the pre-activated language context or **language mode**. In theory, a pre-activated language context enhances the activation level of a language relative to the other, making it easier to process (i.e. language mode (Grosjean, 2001), see also § 4.2.4 Language mode). When bilinguals were experimentally presented with a sentence in one language or the other, they were facilitated in solving the upcoming arithmetic problem when it was presented in the same language as the sentence (Van Rinsveld, Schiltz, Brunner, et al., 2016).

5.3 Model of bilingual number representations

Numerical representations are part of the cognitive processes that are involved in solving arithmetic. Conceptually, to solve an arithmetic problem, numerals need first to be identified and represented before undergoing the cognitive processes underlying the resolution of word problems and arithmetic, likely enacted by the verbal working memory, described in the previous chapter. These representations are activated a second time to retrieve the solution (or only one time if we consider the direct retrieval of arithmetic facts from long-term memory). Empirical data on number transcoding reaction times correlating with arithmetic performances corroborate this idea (Clayton et al., 2020; Steiner, Banfi, et al., 2021; van der Ven et al., 2017).

Verbal numerical representation of numbers can be subdivided into lexical recognition and retrieval and semantic representation of numbers. Lexical recognition and retrieval are the first cognitive processing steps when processing numbers: they form the identity of a numeral. For example when seeing 5 or reading "five", these can be identified as being the same. When reading out loud 5, there is a passage from a visual to a verbal lexical representation. This passage is called transcoding since it involves the passage between two codes of the Triple

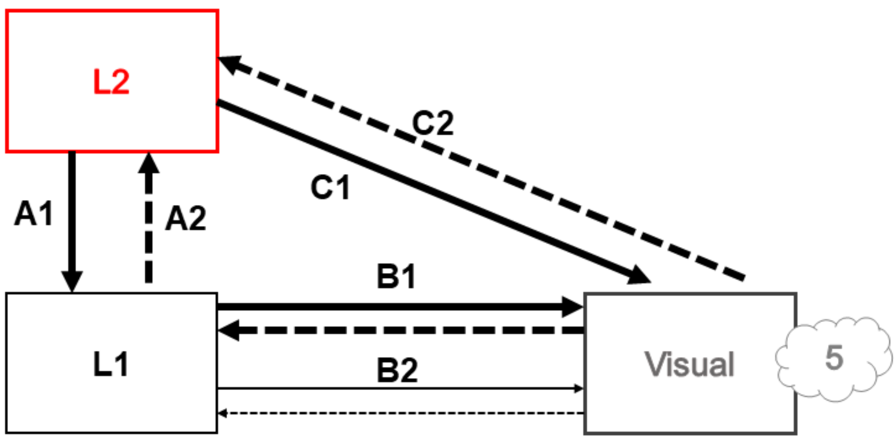
Code Model (see § 2.2 Triple Code Model (TCM). We have already reviewed how language characteristics such as morpho-syntactic transparency affect transcoding in German and French (see § 3.2 Language specific morpho-syntactic transparency), however, we will also see how individual language profiles (§ 4.1 Bilingual heterogeneity in language profiles) can affect how numbers are processed.

5.3.1 Bilingual Encoding Complex Model (BECM)

The Bilingual Encoding Complex Model (**BECM**) has been suggested by (Bernardo, 2001; Garcia et al., 2021) for how bilinguals represent numbers. As the name suggests, it is inspired by the Encoding Complex Model (Campbell, 2005), see also § 2.3 Encoding Complex Model (ECM)). Hence this model suggests numbers are automatically semantically mediated and proposes three different codes for numbers in bilinguals: primary (Arabic digits), secondary (L1) and tertiary (L2). The difference between each code's activation and association is determined by experience with each. Hence the secondary code is not forcefully the first language but the language in which mathematics is learned (LM). Finally, the model predicts asymmetric activations across the codes such that weaker codes activate stronger but stronger ones do not activate weaker codes, see **Figure** *7*.

Figure 7

Bilingual Encoding Complex Model



Note: Adapted by the author. A 1 and A2 are symmetric associations between the L1 and L2 (with stronger associations from L2 to L1 than L1 to L2). B1 and B2 represent bidirectional L1 associations with visual codes. C1 and C2 represent L2 associations with visual codes. Full lines represent stronger associations.

However, the model is limited to arithmetic verification tasks rather than for general processing of arithmetic and numbers: “Therefore, the BECM is best viewed as a model for arithmetic verification and not as a general all-purpose model for arithmetic performance for bilinguals.” (Bernardo, 2001, p. 974). As pointed out in (Garcia et al., 2021) the BECM, hence determines a hierarchy between the different numerical codes.

5.4 Bilingual lexical representations of numbers

Number words serve as the basis for verbal lexical representations of numbers. Consequently, bilinguals compared to monolinguals, have for each number two verbal lexical representations available (i.e. “cinq” and “fünf”). Number words display more variability since they depend on languages which stem from oral transmission, compared to the symbolic written notation of numerals, which are more shared across languages (i.e. the Arabic numeral “5” is shared cross-linguistically, see Chrisomalis, 2010; Ifrah & Bellos, 2000). Concretely it means that for many bilinguals there are two lexical mappings for a single Arabic numeral: “fünf” ↔ 5 ↔ “cinq” (Salillas & Martínez, 2018).

Lexical competition can thus result from these parallel bilingual lexical associations. When a bilingual names “5”, both language’s number words might get activated in parallel, thus requiring a cognitive control mechanism that is not necessary for monolinguals. Another specificity of bilingual number processing with regards to lexical associations is the lexico-lexical associations between number words in multiple languages, for example, bilinguals can translate “cinq” in “fünf”. Given these different types of lexical associations, bilinguals might have different strengths of lexical associations depending on their language status. The level of activation (or resting level) of each number word might differ according to language profiles (i.e. language proficiency, balanced/unbalanced, age of acquisition and exposure, see § 4.1 Bilingual heterogeneity in language profiles) as well as the general (i.e. language mode see § 4.2.4 Language mode) and specific language context (i.e. language of the previous trial or learning mathematics § 5.2 Language Switching Cost (LSC)). Finally, having different lexicons can also result in morpho-syntactic cross-linguistic differences as we have seen in § 3.2 Language specific morpho-syntactic transparency, see also (Dowker & Nuerk, 2016).

The bilingual lexical representation of each number depends on the language profile with respect to the home language (**HL**) and the language for learning mathematics (**LM**). However, also differences between languages can affect lexical access: for example, if a second learned language is less transparent it might complicate learning and might even affect adult's performance. Moreover, bilinguals might already be slower in lexical access in their first language compared to monolinguals. For example, number naming is faster for monolinguals, than bilinguals and even slower for trilinguals (Mägiste, 1979). In the second language, the strength of association depends on various factors such as AoA, language exposure and proficiency (cfr § 4.1 Bilingual heterogeneity in language profiles) hence varying across bilingual profiles. For example, bilinguals who acquire a second language incidentally during school can be described by the following profile: one (or more) home language(s) (HL) and a second acquired schooling language. While the HL improves general language exposure, it might not be the best predictor for the strength of lexical representations of numbers.

5.5 Bilingual Morpho-syntax effect on accessing lexical representations

Regarding **transparency of power**, Lafay et al., (2023), compared Canadian bilinguals with French as LM and either French or another Language as HL in a number writing task. They found a similar quantity of lexical and syntactic errors for numerals in the French base-20 system for both groups with French and groups with another language as HL. Hence suggesting a similar quantity of lexical and syntactic morpho-syntactic errors for the LM, independently of the HL. Using EEG, Salillas & Carreiras, (2014) investigated balanced and equally proficient Spanish-Basque bilinguals, however, half of the participants learned mathematics in Spanish (LM = Spanish) and the other half in Basque (LM = Basque). The participants underwent a number comparison task, where two two-digit Arabic numerals were sequentially presented, and the participant had to judge if the second was larger or smaller than the first. The distance between the numbers was manipulated to have close and far pairs. Furthermore, the pairs were constructed to facilitate or not the number word structure in Basque. Basque number words follow a base-20 morpho-syntactic structure, while Spanish follow a base-10. Hence the Basque pairs were constructed as follows: for example, 40-56 because if 56 is "cincuenta y seis" in Spanish (50 + 6, fifty-six) it is "berrogeita hama sei" in Basque: (2*20 + 16, two twenty sixteen). The ERP results showed an N1-P2 component for the Basque pairs only for the group that learned mathematics in Basque. Hence this study indicated that early cognitive processing of Arabic numerals might be more strongly influenced by the number word structure of the language in which mathematics are learned (LM).

Transparency of order can be observed affecting monolinguals and bilinguals in the compatibility effects in number comparison tasks. Such that responses are slower for 45 < 71 than 45 < 78 because the ten and unit part are both larger (i.e. Nuerk et al., 2002). (Van Rinsveld, Schiltz, Landerl, et al., 2016) compared German(LM1)-French(LM2) bilinguals with language-matched monolinguals on a number comparison task focusing on the compatibility effect (i.e. responses are slower for 45 < 71 than 45 < 78 because the ten and unit parts are both larger). The stimulus pairs varied in terms of compatibility (i.e. if both ten and unit are larger for one numeral than the other as in 54_78) and distances between the tens and units (for example, 54_38 is an incompatible trial with a small distance between tens but a large between units). These were either presented as Arabic digits or auditorily. With Arabic numerals, they found similar effects for the German monolinguals and bilinguals in German and in French (i.e. larger compatibility effect when the distance between units is larger), while the French-speaking monolinguals compatibility effect was not modulated by the unit distance. The same pattern was observed for number words. Xenidou-Dervou et al., (2023) investigated Dutch-English(L2) bilingual participant. In an innovative task, they created artificial number words in Dutch to make them match the quantity and/or congruency of subsequent Arabic numerals. For example for 42, this could be “forty-two” and “two-forty”, matching quantity but the former being incongruent with Dutch ten-unit order. Or, for non-matching “twenty and four” and “ four and twenty” were presented, where only the former is congruent with the Arabic numeral order 42. Dutch(L1)-English bilinguals had to match these numerals and made more errors with the “two-forty”, which is the congruent ten-unit order in Dutch, than for “forty-two”. Furthermore, the quantity of errors additionally decreased for the participants with high proficiency in English. Hence indicating that the (non-inverted) morpho-syntax of the L2 affects the processing of artificial number words built in the L1.

5.6 Bilingual lexico-semantic association of numbers

In the previous chapter, we have seen how bilinguals’ association of symbolic numbers’ lexical representations can differ depending on individual profiles and their access can be influenced by language morpho-syntactic differences. Those differences seem to appear or be more accentuated for some bilingual language profiles such as late sequential learners, in the specific domain of mathematics. At this point, an additional question arises: are these differences limited to lexical retrieval or do they also affect semantic access? In other words, when a slower activation of a number word in one language compared to the other is observed, does it also involve the activation of its meaning?

The difference between the lexical and semantic representation of numbers is that, while the lexical representation concerns the representation of a single numbers, semantic representations concern the quality of association between different numbers. Investigating the question of lexico-semantic associations from the numerical standpoint has several advantages compared to general linguistics since in several different languages there is a common corresponding symbol for numerals (i.e. Arabic numerals). Furthermore, the investigation of the semantics of number words is facilitated compared to other words given that number words have strictly similar semantics across languages. For example "cinq" or "fünf", independently of languages share the same cardinality (i.e. "there are 5 elements"), ordinality (i.e. "it's the 5th element and it comes after the 4th"), magnitude (i.e. "it's 5 times larger than 1"), distance (i.e. "it's at distance 5 from 1 and 7"), parity (i.e. "5 is odd"), etc. However, the strength of lexical semantic associations, the associations between the identity of a numeral and its meaning, might differ among number words in bilinguals.

5.6.1 Cardinality

Bilinguals need to learn to count and become **cardinal knowers** in two languages (see § 3.3.1 Cardinality and Counting). Hence raising the question of whether this knowledge acquired in one language automatically transfers to the other.

Marchand et al., (2020) investigated bilingual English (L1) – French (L2) 5 to 7-year-old children, who were tested in a rapid numerosity of dots naming task in both of their languages. Children were faster at naming dot's numerosities in English than in French even after controlling for word familiarity. Hence, despite this study did not distinguish between L1 and L2, they found independence between language's number knowledge. In the author's terms, these results show that each language is mapped independently with quantity. Wagner et al., (2015) tested 2 to 5-year-old English-Spanish and Spanish-English bilingual children, using a give-a-number task, where children are asked to give a certain number of objects and the dependent variable is the actual number of objects given. The results show that L1 counting proficiency (i.e. which cardinal knower level the children had in one language) did not predict the performance in L2. In other words, when children knew how to use and hence the cardinal knowledge of "three" in English this did not mean they had also acquired the Spanish correspondent "tres". However once the cardinal principle was acquired this could be predicted by both language's counting proficiency. Hence early number representations are acquired separately in each languages, however once the cardinality principle is acquired, this can transfer across languages.

5.6.2 Magnitude size effect

Duyck & Brysbaert, (2004) used a number naming and translation task where balanced or unbalanced Dutch French bilingual participants had to name number words ranging from 1 to 12. Numerals had to be either named in the language they were written in or to be translated. The result showed that larger numbers words were named slower than smaller numbers (i.e. **number magnitude effect**) only when they had to be translated, both in balanced and unbalanced bilinguals. No magnitude effect was found for naming Arabic numerals nor reading number words in L1 or L2. Furthermore, a number magnitude effect could be found for translating learned novel number words (i.e. number words in Estonian). Hence suggesting that translation might be semantically mediated. Also, the author suggests that the semantic connections of number words can build rapidly in development. Duyck and Brysbaert (2008) investigated Dutch (L1)-English(L2) late bilinguals (i.e. AoA for English was 14 years old) in a semantic magnitude task where Arabic numerals, L1 and L2 number words had to be named or translated. Unlike in the previous study from the same author, the task did not lead to a magnitude effect for forward translations (i.e. L1 number words translated in L2), but only for backward translation. The authors explained the different results of the presence of forward translation in Duyck and Brysbaert (2004) but not in Duyck and Brysbaert (2008) by the greater lexical differences in the languages compared in the first study (Dutch vs. French) compared to the second (Dutch vs. English) (Duyck & Brysbaert 2008). For example "zeven" in Dutch, resembles more "seven" in English than "sept" in French. This orthographic overlap, is argued, leads to stronger word-form connections across similar languages. In sum these studies indicate that the semantic magnitude might be activated when translating number words in bilinguals, but that this activation might be diminished if both languages share lexical similarities, thus leading to automatic lexical retrieval as for reading L1 number words.

5.6.3 Priming Number Distance Effect

As we have seen previously (see § 3.3.2 Lexico-semantic associations) **PDE** also works across numerical formats (i.e. dot clouds, Arabic numerals, or number words), it is an interesting paradigm to assess the associations between those formats. It is particularly interesting for bilinguals since it works with number word primes and Arabic numeral's target to be named (i.e. Reynvoet et al., 2002), allowing to test the lexico-semantic associations of several languages (i.e. different number word primes naming languages) in within-subject designs. Moreover, when the primes are short enough, these become inaccessible for consciousness to perceive (i.e. the participants cannot say if or what was primed), meaning that the lexico-semantic associations can be tested implicitly.

Duyck and Brysbaert (2002) investigated Dutch (L1) and French (L2) bilinguals in Belgium with a PDE design. Primes were always Arabic digits presented for 42 ms, Targets were either Arabic digits or number words in the L1 or L2 as targets. The participants were subdivided into two groups and had to either name the targets in the L1 or L2. Therefore some targets had to be translated (i.e. when the French naming group saw a Dutch number word). Their results reveal a PDE distance effect for the target's number word naming and translation: both forward when the Target in L1 is named in L2 (i.e. prime = 2, target = "vijf", response = /vijf/ or /cinq/), and backwards when the Target L2 is named in L1 (i.e. prime = 2, target = "cinq", response = /cinq/ or /vijf/). Hence the authors concluded that there are no differences in the lexico-semantic activations in L1 and L2. And that translating from L2 to L1 (i.e. backwards) was favoured compared to translating from L1 to L2 (i.e. forward).

Duyck et al., (2008) investigated Dutch (L1)-English(L2)-French(L3) trilinguals with a cross-lingual PDE. In this design, L2 English number words were used as primes for all trials (using forward and backward masks). These were followed by Dutch and French number word targets, which, in different blocks had to be named or translated in L1 or L3. Hence leading to two naming and two translation conditions. The results showed that L2 primes facilitated forward and backward translation when they represented the same number (i.e. repetition priming). Results matching with PDE could only be found when naming in L1 and for backward translation of L3 targets to the L1. (i.e. L2 Primes, L1 and L3 Target named in L1). However, no clear PDE could be found for forward translation (i.e. L1 to L3) nor naming in L3. (i.e. L2 Prime, L1 and L3 Target, named in L3). In other words, the results suggested strong L2 lexical associations with translation equivalents (i.e. "five" → "vijf" and "cinq" to be translated into /cinq/ and /vijf/). But lexico-semantic associations could only be found when targets had to be named or translated in L1 but not in L3 (i.e. "five" → "twee" named /twee/ and "deux" to /twee/) and not in the other cases (i.e. i.e. "five" → "deux" named /deux/ and "twee" to /deux/). In short, L2 primes facilitated lexical access both of L1 to L3 and L3 to L1 translations. Lexico-semantic access was however only found for naming in L1 and backward translations.

In sum, these studies suggest lexico-semantic (i.e. measured with PDE) associations between prime number words/Arabic numerals and target number words in both languages. When target number words have to be translated, lexico-semantic associations resulted in being larger for backward than forward translations.

6 Introduction to the studies

In the first chapter we have seen evidence that symbolic numerals are essential for exact representation of numbers. Specifically for an important developmental role of symbolic verbal numerals (i.e. number words). In § 2 Models of numerical representations, we have then reviewed several cognitive models accounting for different representations of numbers and their semantic associations. These models underline the importance of language in verbal representations of numbers (see § 3 Language-dependent influences on the exact verbal representation of numbers). In this chapter we have reviewed three levels language can influence verbal representations of numbers: morpho-syntactic, lexical, and semantic. In § 4 Bilingualism we have seen various factors that constitute the heterogeneity in bilingual's language profiles, followed by models and neuro-cognitive evidence of how bilinguals processes languages. Finally, in § 5 Bilingual numbers we have reviewed how bilingualism affect mathematical performances. We have then seen two cognitive models for how bilinguals' processes numbers. In sum, we have reviewed some evidence for how bilinguals' access to number representation can be differently affected by language's morpho-syntax, lexical and semantic associations.

Given the heterogeneity of bilingual general population, and in a form of ecological control for language acquisition, we carried our investigations on samples from a specific **Luxembourgish population**. This population is homogeneous with regards to formal (i.e. by education) acquisition of numbers, since all the participants of our bilingual samples where following or followed the bilingual education system of Luxembourg. This bilingual education system consists in teaching in German for the first years and then switching to French, see Table 1.

Table 1

Summary of Luxembourgish education system with focus on teaching language

EU education:	Matérnelle	Matérnelle		Primary						Secondary (1 to 7 grades)						
Lux-education:	Kindergarten	Cycle 1 (C1)		Cycle 2 to Cycle 4						Classes (7 to 1)						
US-Grade:	Pre-school	Pre-school		1	2	3	4	5	6	7	8	9	10	11	12	13
Age:	1 to 3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
LM:		-[1]		German						French						
All subjects:	Luxemb.	Luxemb.		German						German			French			

Notes: Luxemb. = Luxembourgish, LM = language of Leaning Mathematics. Only non-vocational track is represented here, vocational differentiation occurs at US grade 10. Age is calculated for starting at 6 years old and excluding class repetitions. [1]pre-schoolmight involves the first steps in numerical acquisitions such as counting. Note that by a new law form 2017 French is introduced and learned in Kindergarten.

With this specific sample of German (LM1) – French (LM2) bilinguals in mind, we aimed at investigating the following research questions about bilingual symbolic number representations, which is summarized in Table 2.

- How does increasing LM2 proficiency affect morpho-syntactic processing and lexical access ?
- How does the morpho-syntactic decade-unit position influences German and French monolinguals and bilinguals ?
- How does lexical and lexico-semantic access compare in bilinguals?

In **Study 1** we measured lexical retrieval with an auditory-visual matching task and a number-naming task. We tested different age groups of bilinguals with increasing proficiency in a second language of mathematical learning (LM2), such as implemented by the Luxembourgish school curriculum. We hypothesized that the slower responses in LM2 than LM1 (i.e. LM2 cost) would decrease with increasing LM1 proficiency and worse performances for opaque base-20 LM2 number words compared to the LM1. We found similar language-related effects independently from increasing proficiencies, suggesting no improvement in LM2 lexical retrieval with increasing proficiency (i.e. an LM2 cost). Moreover, 70' to 90' numerals in French, which are in base-20, resulted in an additional cost for the studied tasks. Hence the morpho-syntactic property of number words (i.e. transparency of power) compared to Arabic numerals influenced number processing.

Study 2 uses an experimental approach to answer the question of how units and ten morpho-syntactic positional order are processed in LM1 and LM2 compared to monolinguals. We used an auditory-visual matching task, where we manipulated the presentation of the ten or unit part of the numerals. This time we only had adult participants but compared bilinguals with language-matched monolinguals. We hypothesized a larger effect of the unit-first condition mimicking German in German monolinguals and bilinguals in German. For French monolinguals and bilinguals in French we hypothesized faster responses for the ten first condition mimicking French number words. For bilinguals, we hypothesized slower responses for bilinguals (bilingual lexical cost) and and LM2 cost (slower in the LM2 than LM1). The results showed that bilinguals in French, not only were slower than in German (i.e. LM2 cost), but they were relatively interfered when cued by the unit part of two-digit numerals but facilitated when cued by the ten parts. Hence these results suggest that morpho-syntactic properties of number words can flexibly over-generalize from the LM1 to LM2 (i.e. interfering in LM2 processing), but that LM2 transparency also facilitates processing relatively to monolinguals. Finally we could confirm the bilingual lexical cost, such that bilinguals in German were slower than monolinguals in German.

In **Study 3** we used a Priming Distance Effect (PDE) design by priming number words in German (LM1) and French (LM2) followed by Arabic numeral targets to be named in both of bilingual's languages. We hypothesized slower response in French than in German (LM2 cost) and PDE with LM1 number words. The results showed that while the German number word prime elicited a PDE, this was not observed with French number words. Hence suggesting that the lexico-semantic weight of associations differs among both languages.

Table 2

Summary of the three studies

Study	Paradigm	Design	Research Question	Group/Sample(s)	Stimuli	Hypothesis
1	Number naming Number matching	Within ID Between Lang.	How does increasing LM2 proficiency affect morpho-syntactic processing and lexical access ?	*Children & adults Bilingual* 26 (5th g); 28 (8th g); 25 (11th g); 20 (adults)	*2-digit* Auditory ↔ Visual (AN)	• Increasingly slower in LM2 • Slower for less transparent numerals. But decreasing with age/proficiency
2	Auditory-visual number matching	Within & between ID and Lang.	How does the morpho-syntactic decade-unit position influences German and French monolinguals and bilinguals ?	*Adult Bilingual & Monolingual* 55 (MonoDE); 56 (MonoFR); 50 (Bilinguals)	*2-digit* Auditory → Visual (AN)	• Slower in LM2 • Bilinguals slower than monolinguals • Flexible morpho-syntax processing in LM2
3	Priming Distance effect (PDE)	Within ID Between Lang.	How does lexical and lexico-semantic access compare in bilinguals?	*Adult Bilinguals* 32 (Bilinguals)	*1-digit* Visual (AD) → verbal (NW)	• LM2 cost • Repetition priming in both languages • PDE only with LM1 primes

Notes. ID = Participants, Lang. = Languages. g. = grades. MonoDE = Monolingual German. MonoFR = Monolingual French. AN = Arabic Numerals. NW = Number Words. LM1/2 = first/second language of math acquisition. Lexical access. Morpho-syntactic influence on lexical access. Lexico-semantic access

1 STUDY 1

Number transcoding in bilinguals - a transversal developmental study

Rémy Lachelin[1], Amandine van Rinsveld[2], Alexandre Poncin1 and Christine Schiltz[1]

[1] Institute of Cognitive Science and Assessment, Department of Behavioural and Cognitive Sciences, University of Luxembourg, Esch-sur-Alzette, Luxembourg

[2] Graduate School of Education, Stanford University, Stanford, CA, United States

Published:

Lachelin, R., Rinsveld, A. van, Poncin, A., & Schiltz, C. (2022). Number transcoding in bilinguals—A transversal developmental study. *PLOS ONE*, *17*(8), e0273391. *https://doi.org/10.1371/journal.pone.0273391*

1.1 Abstract

Number transcoding is the cognitive task of converting between different numerical codes (i.e. visual "42", verbal "forty-two"). Visual symbolic to verbal transcoding and vice versa strongly relies on language proficiency.

We evaluated transcoding of German-French bilinguals from Luxembourg in 5th, 8th, 11th graders and adults. In the Luxembourgish educational system, children acquire mathematics in German (LM1) until the 7th grade, and then the language of learning mathematic switches to French (LM2). French `70s `80s `90s are less transparent than `30s `40s `50s numbers, since they have a base-20 structure, which is not the case in German. Transcoding was evaluated with a reading aloud and a verbal-visual number matching task.

Results of both tasks show a cognitive cost for transcoding numbers having a base-20 structure (i.e. `70s, `80s and `90s), such that response times were slower in all age groups. Furthermore, considering only base-10 numbers (i.e. `30s `40s `50s), it appeared that transcoding in LM2 (French) also entailed a cost. While participants across age groups tended to read numbers slower in LM2, this effect was limited to the youngest age group in the matching task. In addition, participants made more errors when reading LM2 numbers.

In conclusion, we observed an age-independent language effect with numbers having a base-20 structure in French, reflecting their reduced transparency with respect to the decimal system. Moreover, we find an effect of language of math acquisition such that transcoding is less well mastered in LM2. This effect tended to persist until adulthood in the reading aloud task, while in the matching task performance both languages become similar in older adolescents and young adults. This study supports the link between numbers and language, especially highlighting the impact of language on reading numbers aloud from childhood to adulthood.

2 Introduction

Exact numerical representations are supported by symbolic verbal (*e.g.* forty-two) and visual (*e.g.* 42) representations which are acquired through learning. However, in many languages the verbal number word's syntax differs from the visual one, *i.e.* the Arabic number system (Chrisomalis, 2010; Ifrah & Bellos, 2000). This difference might have an influence on transcoding, *i.e.* the cognitive transformation from one code to another. Thus reading aloud for instance the number "42" implies the transcoding from a visual ("42") to a verbal code ("forty-two"). Moreover, in the case of bi- and multilingualism, acquiring multiple languages means that during development multiple verbal codes are mapped with the visual code (Salillas & Martínez, 2018). Since multilinguals are outnumbering monolinguals across the globe (Grosjean, 2010), the question of how numbers are processed and particularly transcoded using two (or more) verbal codes in a (developing) multilingual cognitive system is of crucial importance (Wicha et al., 2018).

Several cognitive models have been proposed to describe transcoding. While some are taking also into account cognitive development, most of these models do not specifically account for transcoding differences between languages and multilingualism. These models are summarized below in two main categories: semantic and asemantic (Barrouillet et al., 2004).

In semantic models, transcoding requires an obligatory access of the number's magnitude. For example McCloskey (McCloskey, 1992; McCloskey et al., 1985) proposed a transcoding model in which the entry number - regardless from the input's format - accesses an abstract semantic representation. Power and Dal Martello (Power & Dal Martello, 1990) further proposed a model specifically for number dictation (*i.e.* verbal to Arabic format) which differs from the previous one in that semantic representations reflect the verbal word structure of the numbers, thus predicting differences between languages. Semantic models hence assume that transcoding difficulties depend on the quality and maturity of the semantic representations, therefore predicting worse performances in children, in particular for larger numbers.

Asemantic models propose that numbers can be transcoded without accessing magnitude. Deloche and Seron (Deloche & Seron, 1982) proposed such a model for number naming (*i.e.* Arabic to verbal) where the first step involves parsing the input into primitives which are then submitted to a set of rules for the output system. Later on, Barrouillet and colleagues (Barrouillet et al., 2004) proposed ADAPT (A Developmental Asemantic Procedural Model), explicitly including a developmental perspective on transcoding. In ADAPT, by repeating transcoding, number words become lexicalized with training, hence directly retrieved from long-term memory and bypassing the procedural processes. Therefore transcoding in children would depend on language-dependent characteristics. Yet, the language-dependent characteristics should diminish with increasing lexicalization. Another more general model considering number representations is the Triple Code Model (TCM). The TCM proposes a functional and topographical framework of how and where in the brain approximate, visual symbolic and verbal codes are processed. The TCM also implies asemantic language-dependent transcoding from or to a verbal code (Dehaene, 1992). Language-specific training in transcoding would therefore increase the strength of the association between each code and their respective brain areas. More recently, Dotan and Friedmann (Dotan & Friedmann, 2018) proposed a model for Arabic to verbal transcoding where numbers are first visually analysed for identity, order and decimal structure to build a language independent number word frame, which is passed to a second system which applies language-specific rules associated with their phonological and articulatory counter-parts. In sum, asemantic models predict that verbal transcoding depends on language characteristics. However those models do not explicitly model multilingualism, nor how transcoding develops by acquiring number words in a second language.

When investigating how the acquisition of several languages (and hence multiple number words) develops, the many forms and constellations of multilingualism must be taken into account; with bilingualism corresponding to the simplest instance. Today, multilingualism is often considered as a continuum that is shaped by numerous aspects such as relative language proficiency (de Groot, 2011; Weber-Fox & Neville, 1996), as well as age and duration of language learning (Fiebach et al., 2003). Other factors influencing multilingual's profiles also depend on years of education (particularly literacy education), the amount of language input and output, and its privileged context (*e.g.* (Grosjean, 2001)). Thus, multilinguals often have a first language (L1) which is more dominant than later learned languages (*i.e.* L2). On top of these factors, the structure of a language such as morphemes and syntax shape each language learning trajectory. These factors might therefore impact transcoding as well as arithmetic in general, since common processes are implicated in both (Clayton et al., 2020; Steiner, Banfi, et al., 2021; van der Ven et al., 2017). See also (Dowker & Nuerk, 2016) for an overview of language-related factors. In the following, we will briefly review two main factors that are critical with regards to our study: *Transparency of power* and *language of math acquisition*.

2.1 Transparency of power

Transparency of power refers to the existence of different degrees of morpho-syntactic language transparencies in verbal numbers. In many Asian languages (*i.e.* Mandarin Chinese, Vietnamese) the morpho-syntactic structure of the verbal number system is highly consistent with the Arabic number system, and in general with the base-10 (*i.e.* 10*x + y, where x indexes the base and y the unit, see (Miura et al., 1988)). This linguistic characteristic, also termed *transparency of power*, facilitates learning to count (M.-L. T. Lê & Noël, 2020; Miller et al., 1995; Miller & Stigler, 1987) and solving arithmetic problems (Geary et al., 1996; McClung & Arya, 2018; Rodic et al., 2015). Another morpho-syntactic difference concerns the inversion between tens and ones as for example in Dutch and in German (*i.e.* y + 10*x; in German 42 is "zwei-und-*vierzig*", literally "two-and-*forty*"). This linguistic characteristic, called *transparency of order* (Bahnmueller et al., 2018), explains some transcoding error patterns and reaction times (Clayton et al., 2020; Imbo et al., 2014; Moeller, Zuber, et al., 2015; Pixner et al., 2011; Poncin et al., 2019; Steiner, Banfi, et al., 2021; Zuber et al., 2009), gives rise to specific pattern of compatibility effects (Bahnmueller et al., 2015; Nuerk et al., 2004, 2005) and complicates the solution of certain arithmetic problems (Göbel, Moeller, et al., 2014; Lonnemann & Yan, 2015; Xenidou-Dervou et al., 2015).

Transparency of power can also vary with a change of base, for example a base-20 system (20*x + y), also called vigesimal. The use of the vigesimal system is found for example in French, Basque, or in Diola-Fogny (see (Haspelmath et al., 2005)). French, to be precise, has a mixed system since it uses base-10 and base-20 systems. Up to the `60s in French, like in English, the counting system follows the base-10 rule. However, `80s to `90s follow a base-20 system, e.g. 87 = 4*20 + 7, "quatre-vingt-sept", literally "four-twenty-seven". Moreover, the teens (11 to 16) are lexical primitives, like in English and in stark contrast to the more transparent Asian languages. Note that the transparency contrast is additionally increased for 71 to 76 and 91 to 96, which are composed with a base-20 and teens, e.g. 96 = 4*20 + 16, "quatre-vingt-seize", literally "four-twenty-sixteen". Furthermore, the vigesimal system is subject to regional variances, for example in Belgian Wallonia `70s and '90s are not vigesimal (i.e. "septante" for "seventy" and "nonante" for "ninety", see (Seron & Fayol, 1994)) and in some Swiss-French cantons the vigesimal system is entirely absent (i.e. including for 80, "huitante", literally "eighty"). In a study comparing French-speaking 1st graders from Wallonia and France with a number dictation task, more mistakes were committed in the `70s and `90s by the latter (Seron & Fayol, 1994). In another study comparing English-speaking to French-speaking 5th graders, numbers above 60 were slower to transcode, revealing a cost for base-20 numbers (Van Rinsveld & Schiltz, 2016). Further indicating an interaction between the number structure and number processing, Basque-speaking adults solved additions faster when the operand was in the base-20 structure. For example, the solutions for 20 + 15 is facilitated over 25 + 10, since the result 35, is said “hogeita-hamabost”, literally “twenty-fifteen” (àngels Colomé et al., 2010). Strikingly, for Basque-speakers the distance effect also leads to different event-related potential brain responses for vigesimal compared to decimal Arabic digits (Salillas & Carreiras, 2014). While providing interesting insights into the neuro-cognitive mechanisms of vigesimal number processing, these studies might be limited by potential cultural and educational confounds associated with group comparisons (Geary et al., 1996) or curricular differences (Krinzinger et al., 2011). Hence, between-subject comparisons should be complemented by within-subject designs with multilingual participants. Furthermore, they do not provide information about the developmental trajectories of the acquisition of a language with a different number system since they focus on single age groups. They consequently need to be completed by designs comparing different age groups of multilinguals.

To sum up, the transparency of power refers to the morpho-syntactic structure of a language’s number word system. Results with different populations and using various tasks consistently indicate an advantage for processing numbers in more transparent languages. While some studies started to explore the impact of vigesimal number structures, only a few studies focused on within-subject designs with bilinguals. And to the best of our knowledge, studies on the developmental trajectory in bilingual learners are still entirely missing.

2.2 Language of mathematics acquisition

In a questionnaire asking multilinguals which language they preferentially use for mental calculations, the majority reported a preference for their first language L1, supposedly also their language of math acquisition (LM+) (Dewaele, 2007). In contrast, solving arithmetic problems in the non-preferred language resulted in cognitive costs for vocal answers to arithmetic problems presented as Arabic digits (Marsh & Maki, 1976; Van Rinsveld, Schiltz, Brunner, et al., 2016), auditory (McClain & Huang, 1982), or written number words (Frenck-Mestre & Vaid, 1993). While highlighting the impact of language dominance when doing math, these designs confound general L1 benefits with potential domain-specific benefits from the language in which mathematic was first learned (Spelke & Tsivkin, 2001a).

Training experiments with bilinguals doing arithmetic in both languages indeed indicate a benefit for solving arithmetic problems in the language in which they were trained. Spelke and Tviskin (Spelke & Tsivkin, 2001b) investigated Russian-English bilinguals, who were trained to solve arithmetic either in the dominant language, L1 (Russian) or their L2 (English). The results showed a cost for solving arithmetical problems when switching to the untrained language, independently if the testing was in the L1 or L2, indicating a Language Switching Cost (LSC). LSC in the context of math training was replicated in 9th and 11th graders attending German-French bilingual education curricula, who were trained to solve arithmetic in German or French (Saalbach et al., 2013). Similar LSC was found in German-French (Volmer et al., 2018) and German-English bilingual adults (Hahn et al., 2017). Hence, independently from language dominance, arithmetic and mathematical problem solving are facilitated when tested in the same language as they are learned, and they are accompanied by a cost in the untrained language. These findings underline the importance of how multilingual education school curricula are designed.

The LSC generalizes to more ecological learning settings, showing that mathematical problems are solved more accurately in LM+, even if it is not the dominant language (*i.e.* the LM+ not being the same language as the L1). For example, Bernardo (Bernardo, 2001) investigated arithmetic among high school Filipino-English bilinguals who have Filipino as L1 but specifically learned mathematics in English (LM+). The results indicate a cost for arithmetical problems written in number words in Filipino (*i.e.* being the L1 but LM-) compared to English (*i.e.* being an L2 but LM+), which in turn showed comparable results than to problems presented as Arabic digits. The critical role of LM+ or LM1 is also confirmed by studies on highly proficient bilinguals. Note that here we make a distinction between the LM+, defined as the (only) language of learning mathematics, and LM1 defined as the first language of learning mathematics (which is followed by other languages used later in the learning process). For example in school curricula where the first language for learning mathematics is German (LM1) and math classes later switch to a second language (French, LM2), systematic costs have been found for LM2 arithmetic problem solving despite an equal number of years training the LM2 (Van Rinsveld et al., 2015). A recent meta-analysis found an advantage for solving arithmetic and naming numbers (but not for magnitude comparison tasks) in the L1 (Garcia et al., 2021).

These behavioural findings are confirmed by recent neuroimaging studies on arithmetic and number comparison. Recording electroencephalogram during a true or false judgment of simple multiplications, Salillas and Wicha (Salillas & Wicha, 2012) studied fluent Spanish-

English bilingual adults who had learned arithmetic in either English or Spanish, the LM+ respectively. The problems presented in the LM+ were solved faster and the corresponding event-related potentials showed a larger N400 response than for problems in LM-, independently of the language representing LM+ and LM-. Assessing two-digit number comparison in Spanish-Basque bilinguals (Salillas & Carreiras, 2014), could even show that the numerical distance effect was modulated by the language of math acquisition. Functional magnetic resonance imaging studies revealed that the LM1 recruited more temporal regions, supposedly related to direct semantic retrieval, than the LM2 for simple additions. In turn, the LM2 recruited a network of regions indicating the need for more generic cognitive resources (Van Rinsveld et al., 2017). On the contrary, Cerda et al., (Cerda et al., 2019) recently investigated Spanish-English bilingual children's performance in a multiplication verification task and observed similar ERP responses in both of their languages. Although the language of math acquisition largely impacts bilingual's arithmetic skills, under certain conditions, such as early learning stages, the bilingual brain might consequently reveal some flexibility.

These results tend to support the above-mentioned TCM stating that precise numbers are encoded in a language-dependent format (Dehaene et al., 1999). They fit less with the classical view from bilingualism research stipulating that representations are independent of languages (*e.g.* (À. Colomé, 2001; Van Assche et al., 2012)) and that both bilingual's languages are active, even in situations when only one language is needed (*e.g.* (Perani & Abutalebi, 2005; van Heuven et al., 1998)). However, they agree with recent reports from the bilingualism literature that academic knowledge acquired in one language is retrieved more efficiently in the learning language compared to another language (Volmer et al., 2018). Further research on numerical cognition is consequently needed to fully understand this complex bilingual situation. Studying how language of math acquisition influences number transcoding is especially interesting since the cognitive mechanisms of transcoding are closely related to word retrieval. Yet, such studies are still missing, both in adults and in children.

2.3 Present study

The present study aimed to better understand the interaction between language and numbers by investigating number transcoding of German-French bilinguals from four different age groups. We targeted 5th, 8th and 11th graders, as well as adults to assess performances before and after the switch in the language of mathematical learning implemented in 7th grade in the Luxembourgish education system.

In the Luxembourgish multilingual school system, pre-schools (3 to 5 y.o.) are in Luxembourgish, which is linguistically close to German. The teaching language in primary school (1st to 6th grade, 6 to 12 y.o.) is German, except for French lessons. From 7th grade to 10th grade, the majority of subjects are taught in German, except for mathematics and French lessons, which are taught in French. Then from 11th grade until the end of obligatory school, all topics are thought in French. This multilingual education system aims to render students highly proficient in both German and French. In sum, critically for number transcoding, the language for teaching and learning mathematics switches from German to French at 7th grade (*Ministère de l'Éducation Nationale*, 2022). Hence students' first language of math acquisition (LM1) is German, while their second language of math learning (LM2) is French. Therefore allowing within-subject designs from a sample with the same educational background. Furthermore, German-French bilingualism is interesting concerning number word structures since both languages differ in *transparency of order* and their *transparency of power*. Previous studies on this specific population have reported language effects in magnitude comparison tasks, showing comparable compatibility effects to monolingual German (Van Rinsveld, Schiltz, Landerl, et al., 2016). While arithmetic problems were solved faster in German than in French (Van Rinsveld, Schiltz, Brunner, et al., 2016), at least for complex additions (Van Rinsveld et al., 2015). Since arithmetic correlates with transcoding (Steiner, Banfi, et al., 2021), we expected similar findings with transcoding tasks. The present study investigates the question of the role language proficiency has on number transcoding. Using a cross-sectional design allowed us to study whether and how (a) number word *transparency of power* and (b) *language of math acquisition* influence two-digit number transcoding at different stages of bilingualism.

We explored the effect of *transparency of power* by comparing transcoding of French numbers above 70 (`70s `80s `90s) and below 60 (`30s `40s `50s), following respectively a base-20 and a base-10 structure. We expected that transcoding performances in both bilinguals' languages would improve with age. We hypothesized that independently of bilinguals' age, language would influence number processing such that non vigesimal French numbers (following a base-10 structure) would be transcoded better than vigesimal numbers over 70 (following a base-20 structure), revealing an effect of *transparency of power*.

For the impact of *language of math acquisition* we compared the performances in both of the bilinguals' languages in the four age groups. Transcoding requires to access and retrieve lexical information on numbers stored in long-term memory. Therefore, we predicted that transcoding would be better in German (LM1) than French (LM2). To capture the different facets of this retrieval process (Vander Beken & Brysbaert, 2018), we deployed two complementary transcoding tasks. In the *reading aloud task* participants had to name Arabic digits, while the *verbal-visual matching task* required the matching of spoken number words with the corresponding Arabic digit. Both tasks are assumed to tap into distinct retrieval mechanisms, *i.e.* free recall and recognition respectively. The two tasks might thus reveal a somewhat different result pattern such as more marked linguistic influences in the reading aloud free recall task than on the verbal-visual recognition task, as already observed in Van Rinsveld et al. (Van Rinsveld, Schiltz, Landerl, et al., 2016).

3 Method

3.1 Participants

From the initial full sample of 125, we first excluded participants who reported having French as their mother tongue, as these participants might have acquired French number words outside the context of formal schooling. This led to the exclusion of 25 participants. One additional adult was removed because of an otherwise missing measure in one of the crossed factors. Therefore we excluded a total of 26 participants, leading to a final sample of N = 99. This final sample consisted of four age groups: 5^{th} graders n = 26, age = 10.69*(0.55)*, 8th graders n = 28, age = 13.46*(0.56)*, 11^{th} graders n = 25, age = 16.48*(0.59)*, adults n = 20, age = 23.58*(5.12)*. The language profiles varied among the participants, 66 reported Luxembourgish as mother tongue, 5 Portuguese, 4 German, 4 English, 4 Italian, 3 Polish, and the 13 remaining all reported a different mother tongue.

Despite the different language backgrounds, all participants were currently living in Luxembourg, where Luxembourgish, German and French are official languages that can be encountered in daily life. Importantly the language support for the formal acquisition of mathematics through school curricula is first in German (LM1) from 1^{st} to 6^{th} grade and switches to French (LM2) in 7^{th} grade. Therefore young pupils start the school curriculum with different degrees of proficiency in French and German to progressively become proficient bilingual adults throughout their education. The four age groups of German-French bilinguals were therefore sampled at different ages of LM2 (French) acquisition: 5^{th} graders being before the language switch in mathematics, while the three older groups have gradually increasing experience and familiarity with practicing mathematics in French.

All pupils were enrolled and tested in schools, while the adults were enrolled and tested at the university. The Review Panel of the University of Luxembourg revised and approved the study. Informed written consent was obtained from all the children's parents and adult participants. Monetary compensation after study completion was given to both pupils and adults.

3.2 Material and stimuli

Headsets connected to the Iolab USB Button Box were used to both (a) record the voice onsets of the reading aloud task and (b) play the auditory stimuli for the verbal-visual matching task. The experiments were run with Psyscope XB7 (J. Cohen et al., 1993) on an Apple 13' MacBook. Reaction times were measured with the Iolab USB Button Box from the end of the auditory recording until button press. Stimuli for each task were 28 different two-digit numbers ranging from 31 to 98. These were further subdivided into two sets of 14 distinct stimuli each, one for the German and the other for the French blocks. The experiment was part of a larger study, which lasted 45 minutes, at the end of which the participants were compensated with a gift voucher. Ties (*e.g.* 44) and tens (*e.g.* 30) were excluded from the sets. The list of all the stimuli used can be found in S3 Table 1. The order in which the sets were presented was randomized. In the reading aloud task, the Arabic digits were presented visually on a computer screen. They appeared in the centre of a white screen in black (Arial, font size 90) and remained visible until participant responses. For the verbal-visual matching task, German and French recordings of two female native German and French speakers were presented as auditory stimuli (as in (Van Rinsveld & Schiltz, 2016)). In the verbal-visual matching task, there were four possible visual Arabic numbers: one target and three distractors, which positions varied randomly. The three distractor-stimuli consisted of: *unit distractors* in which one unit was randomly added or subtracted to the heard number (*e.g.* for 42 distractors were 43 or 41). *Ten distractors* in which a ten unit was randomly added or subtracted to the heard number (*e.g.* for 42 the distractor was 32 or 52). *Unit and ten distractors* in which 11 were randomly added or subtracted from the target number.

3.3 Procedure

All participants were tested individually in a quiet room, children in the schools, and adults at the University of Luxembourg. The participants passed both the reading aloud and verbal-visual matching tasks first in German or French in a counter-balanced order. Both tasks started with a five-stimuli training followed by 14 test stimuli. Participants were instructed to respond as accurately and fast as possible.

In the reading aloud task after the participant named the visually presented number, the answers were written down by the experimenter, who started the following trial by button press. Reaction times were measured with voice key from stimuli screen presentation until first vocal onset. In the verbal-visual matching task, the auditory stimuli were first presented via the headsets. Then, four numbers were visually presented on the screen. Reaction time recording started when the auditory stimulus presentation was completed and ended when one of the response buttons was pressed. Stimuli were presented with an inter-trial interval of 500 ms. At the end of the experiment the participants were compensated with a gift voucher.

3.4 Design and statistical analyses

The transcoding of two-digit numbers was compared in a transversal developmental design by Linear Mixed Models (LMM) on the dependent variables reaction time and correct responses. To control for age related variabilities analogous analyses were conducted on *z-score* transformed RT, separately for each group. Furthermore, to control for word length (see (N. C. Ellis & Hennelly, 1980)), we replicated the same analyses taking into account only numbers where the corresponding number words were constituted of four syllables (see S3 Table 1 for which stimuli were included in these analyses). Data were analysed with R (R Core Team, 2013) using the *afex* package (Singmann et al., 2020), while graphs were drawn with *ggplot2* (Wickham, 2016).

4 Results

For the present analyses, we grouped the stimuli in two categories according to their size and number word structure in French: Small numbers (’30s, ’40s and ’50s) and Large numbers (’70s, ’80s and ’90s) in a factor referred to as “Number Size”. That is, ’60s numbers were excluded in order to create two equally sized groups clearly differing with respect to their French number word structure (*i.e.* decimal vs. vigesimal). However, before removing the ‘60s we applied common exclusion criteria to the data of both tasks. First, invalid trials of the reading aloud task including voice key recording errors, misspellings, inversions and confusions were removed, 4.29 % of the initial sample. Second, to exclude aberrant reaction times (RT) we filtered longer than 5 seconds responses leading to the exclusion of 1.09 % trials in the reading aloud task and 1.52 % in verbal-visual matching task. Third, to exclude outlier responses, faster or slower than 3 standard deviations within each individual mean were removed, additionally, thus excluding 1.29 % and 1.52 % of the trials from the initial responses. Then we removed the '60s numbers (accounting for additional 13.71% and 14.07% respectively for both tasks). For reaction time analyses all incorrect responses were removed, accounting for 2.63 % of the reading task and 5.99 % of the matching task, corresponding to the total error rate. As suggested by an anonymous reviewer, we additionally removed the trials following an error which might be affected by post-error slowing in particular given the high accuracy rate (W. Notebaert et al., 2009), thus additionally accounting for 1.95 % and 5.63 % of the trials respectively.

4.1 Analyses

Both tasks were analysed with Linear Mixed Models (LMM) in R using the mixed function from the *afex* package's *mixed* function (Singmann et al., 2020), which relies on *lme4* (Bates et al., 2015). Follow-up analyses were computed with *emmeans* (Lenth, 2021). When our initially designed full model failed to converge, we reduced the model complexity by gradually simplifying the random effect structure until convergence was reached.

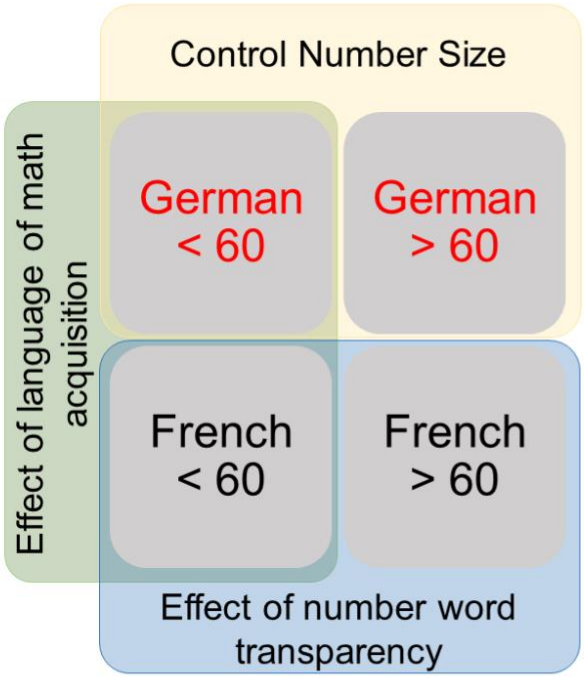


Fig.1 Venn diagram of the design rationale. Design rationale of the comparisons on the different hypotheses of number word transparency (A) and age of math acquisition (B).

The analyses aimed at testing for all age groups the two main hypotheses: (1) the effect of *transparency of power* in French and (2) the effect of *language of math acquisition*. First, the effect of *transparency of power* in French, corresponding to the cost for vigesimal numbers in the base-20 system was tested by comparing small numbers (*i.e.* ’30s, ’40s, and ’50s) to large numbers (*i.e.* ’70s, ’80s, and ’90s) in French only (see A in Fig 1). For this hypothesis, we predicted that both tasks would lead to worse performances, henceforth a cost, for large compared to small numbers. Developmentally, the cost might either be re-absorbed with increasing training using LM2 number words or remain stable across age groups. Additional control for the potential confounder of a number size effect was implemented by applying the same comparison within German (A.1 in Fig 1). Second, the hypothesis concerning the *language of math acquisition* was tested by comparing only small numbers (*i.e.* ’30s, ’40s, and ’50s) in French and German (B in Fig 1). We predicted that both tasks would lead to a cost in French compared to German. Developmentally, this cost could remain stable throughout adolescence and adulthood, or gradually diminish with age, potentially even resorbing in adults.

4.2 Reading aloud task

4.2.1 Reaction Times

The maximal linear mixed model was defined with main effects and interactions between the fixed between-group factor *Age* (levels: 5th, 8th, 11th grade and adults) and two fixed within-group factors: *Language* (German, French), and *Number Size* (Large, Small), all levels being defined as categorical. As random factors, we modelled individual differences (*i.e. Subject*) and item-related variability (*i.e. Item*). The maximum model was defined taking into account individual differences by using random slopes and intercepts *per Subject* for the interaction between *Language* and *Number Size*. Moreover, item-related variability was modelled using random intercepts and slopes *per Item* for the interaction between *Language* and *Age*. This led to the model with the following R syntax form (A0):

$$(A0) \quad RT \sim 1 + Age * Language * Number\ Size + (1 + Language * Number\ Size | Subject) + (1 + Language * Age | Item).$$

However, since the maximal model led to a singular fit (due to high correlations between the random parts of Number Size per Subject and Age per Item) we reduced the complexity of the model by removing those problematic random terms (Barr, 2013). Therefore, the final model takes the following R syntax form (A):

$$(A) \quad RT \sim 1 + Age * Language * Number\ Size + (1 + Language | Subject) + (1 + Language | Item).$$

P-values and degrees of freedom were obtained with the Satterthwaite approximation method by comparing the full model against the model without the effect (Singmann et al., 2020). Follow-ups were calculated by comparing estimated marginal means (EMM) obtained with the emmeans package (Lenth, 2021), and p-values were adjusted with the Bonferroni method. All contrasts were set to sum.

Table 1. Results of the reading aloud task's RT's linear mixed model

	num *df*	**den *df***	***F***	***p-value***
Age	3	92.38	16.43	<0.001
Language	1	97.42	131.18	<0.001
Number Size	1	22.10	39.02	<0.001
Age x Language	3	89.32	16.35	<0.001
Age x Number Size	3	1837.97	3.24	0.02
Language x Number Size	1	21.95	52.82	<0.001
Age x Language x Number Size	3	1838.02	2.92	0.03

Note: degrees of freedom (df) calculated with Satterthwaite approximation; num: numerator, den: denominator.

The model for the reading aloud task RTs resulted in three significant main effects and three two-way interactions (see Table 1). The main effect of Age was decomposed with follow-up analyses showing that the 5th graders (978*(555)* ms) were slower than the older groups, which had similar RTs (8th 788*(282)* ms, 11th 777*(246)* ms and adults 822*(320)* ms; standard deviations in parenthesis). Furthermore, the main effect of Language indicated overall slower naming in French than in German, and the main effect of Number Size indicated slower naming for bigger than smaller numbers. Significant two-way interactions were found between Age and Language, Age and Number Size, as well as Language and Number Size. Finally the three-way interaction was also significant, see Table 1 and Fig 2.

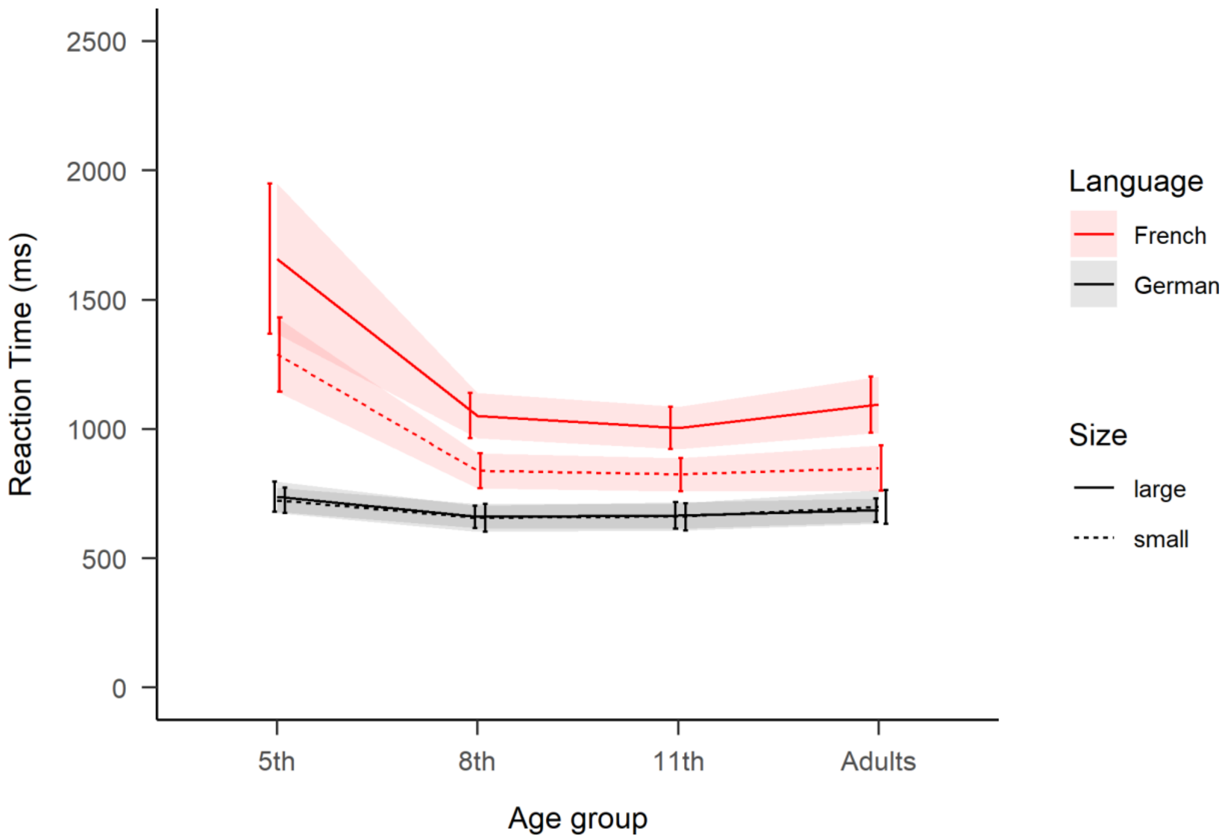


Fig 2. Reaction times of the reading aloud task. For each age group as a function of "Language" and "Number Size". Large numbers correspond to '70s, '80s and '90s, while small numbers correspond to '30s, '40s and '50s. Ribbons represent standard error.

Follow-up analyses were performed on the model's estimated marginal means (EMMs) of the interaction's terms (see Singmann, 2021) with Satterthwaite estimation of degrees of freedom. Then EMM were compared with two-tailed pairwise tests with Bonferroni adjusted p-values. This confirmed the first hypothesis on number word transparency (cf. A in Fig 1), since in French the vigesimal '70s, '80s and '90s numbers were named slower than '30s, '40s and '50s numbers by all age groups: 282 ms slower for 5th graders, 219 ms for 8th, 188 ms or 11th and 236 ms for adults (all $t > 5.13, p < .001$). While as control (cf. A.1 in Fig 1), there was no significant difference in German (max difference 9 ms) between naming large or small numbers (all $t < 0.30, p = 1.0$.).

Secondly, the hypothesis on *language of math acquisition* was confirmed by the two youngest age groups (*cf.* B in Fig 1). Naming in French (LM2) compared to German (LM1) was significantly slower by 510 ms for 5th graders($t(126.17) = 9.09, p < .001$) and by 177 ms for 8th graders($t(117.4) = 2.83, p < .05$). Although positive, the difference was not significant for the two older age groups: 161 ms for 11th grades ($t(117.9) = 2.42, p = .10$)

and 158 ms for adults ($t(117.17) = 1.94, p = .32$). To further establish the robustness of these results, we replicated the analyses using the same model (A), but on RTs transformed into z-scores per age-group (see Fig 3.) and on a sub-sampled dataset including only four-syllable long number words (see S2).

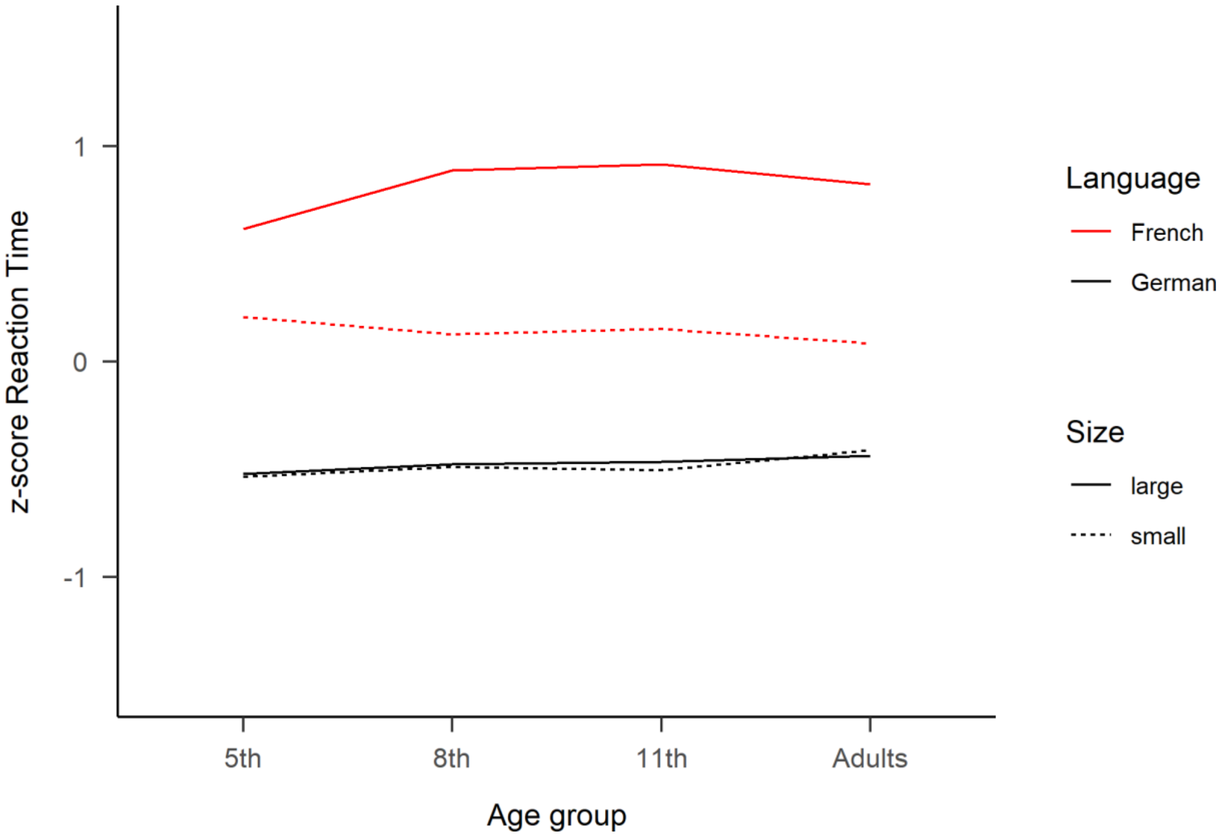


Fig 3. Z-score reaction times of the reading aloud task. Standardized reaction times for each age group as a function of "Language" and "Number Size". Large numbers correspond to '70s, '80s and '90s, while small numbers correspond to '30s, '40s and '50s.

The z-score transformation aimed to reduce the variability of RT among age groups and confirmed the main effects and the interaction between Language and Number Size, while the triple interaction was not significant anymore (see S2 Table 1). Using the dataset with number words of 4 syllables (see S2 Table 3) to control for the possible word length effect confounder ((see N. C. Ellis & Hennelly, 1980)) also replicated the main effects and the two-way interactions, but again the three-way interactions between Age, Language and Number Size was not significant anymore.

4.2.2 Correct Responses

Correct responses (CR) were analysed using a binomial approach and *p-values* estimated by the likelihood ratio test. Since applying the same model as for RT (see (A)) did not converge, we had to drop the random per-*Item* factor and the fixed factor *Age*. In sum the final model had the following syntax (B):

$(B)\ CR \sim Language * Size\ +\ (1|Sujet)$

The failure of convergence with the more complex model might be due to ceiling effects, as the task was very simple, particularly for older age groups, see Fig 4 and S1 Table 2.

Table 2. Results of the reading aloud task's CR's linear mixed model

	df	χ^2	*p-value*
Language	1	47.31	<0.001
Number Size	1	0.62	0.43
Language x Number Size	1	3.90	0.05

The binomial model (see (B) and Table 2) indicated a significant main effect of Language and in contradiction to the RT analyses, the main effect of Number Size was not significant. However, critically, like for RT, the interaction between Language and Number Size persisted with CR.

Follow-up analyses confirmed with CR the pattern observed with RT: in French, 4.8% more errors were made with '70s, '80s and '90s than '30s, '40s and '50s numbers ($z = -3.37, p < .01$), indicating a cost for vigesimal number transcoding. The same contrast in German did not lead to any significant differences ($z = 0.63,\ n.s.$). Secondly, comparing '30s, '40s and '50s numbers in both languages revealed 2.6 % more errors were made in French than in German ($z = -2.89,\ p < .05$), pointing towards a disadvantage for transcoding numbers in French (LM2).

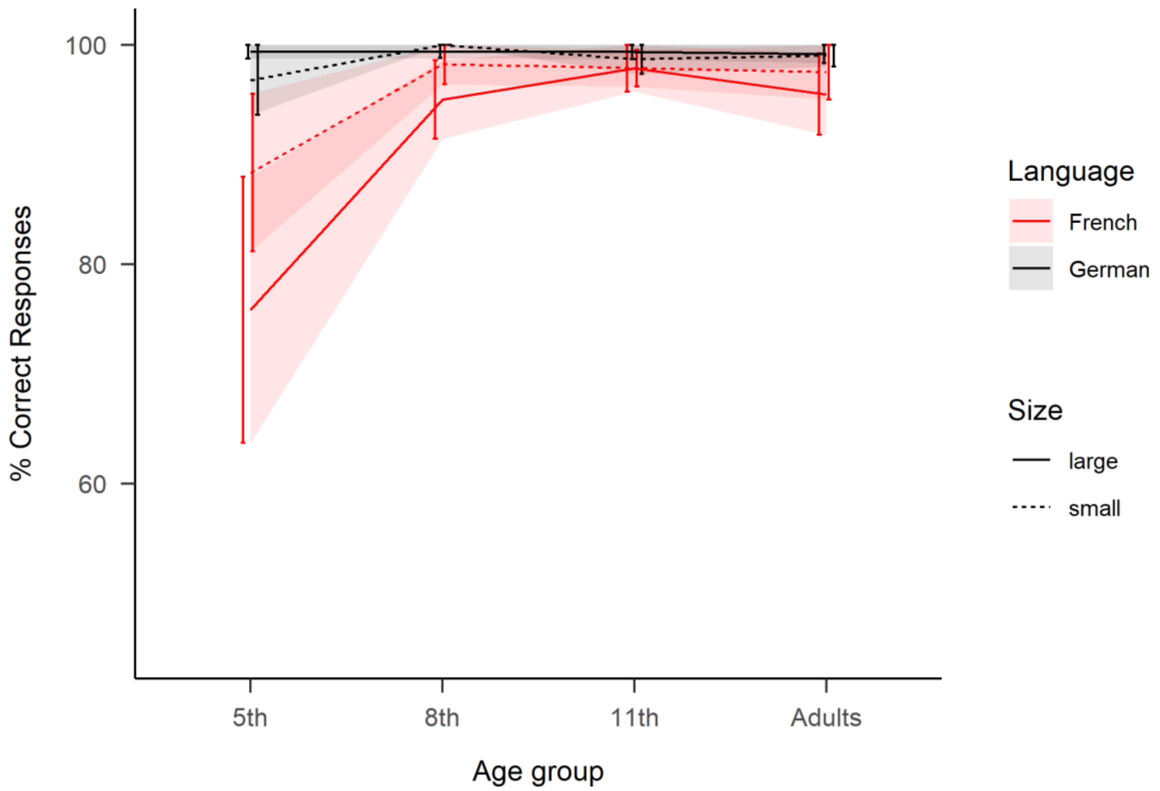


Fig 4. **Correct Responses in the reading aloud task.** Percent correct for each age group as a function of "Language" and "Number Size". Large numbers correspond to '70s, '80s and '90s, while small numbers correspond to '30s, '40s and '50s. Ribbons represent standard error.

In summary, the analyses on RT and CR confirmed both hypotheses which predicted effects of (1) *transparency of power* penalizing vigesimal number words and (2) *language of math acquisition* benefitting LM1 (*i.e.* German). The effect of *transparency of power* was robust across age groups since it could be replicated on z-score transformed data and by limiting the analyses to four-syllable long number words. The effect of *language of math acquisition* was found only in the two youngest age groups of the initial analyses, but persisted across age groups in the two additional analyses.

4.3 Verbal-visual matching task

4.3.1 Reaction Times

For coherence of interpretation and comparison with the previous verbal-visual matching task, the same model (A) and parameters for p-values, degrees of freedom and follow-up's were applied here on RT. That is the factors Language, Age and Number Size were modelled as fixed effects. Additionally the model included random factors to consider item-related variability as a function of language and random intercepts to account for individual differences per subject.

Table 3. Results of the verbal-visual matching task's RT's linear mixed model

	num *df*	**den *df***	***F***	***p-value***
Age	3	92.13	49.90	<.001
Language	1	78.27	76.63	<.001
Number Size	1	21.32	29.00	<.001
Age x Language	3	83.25	14.64	<.001
Age x Number Size	3	1731.94	15.06	<.001
Language x Number Size	1	22.31	78.61	<.001
Age x Language x Number Size	3	1730.32	11.74	<.001

Note: degrees of freedom (df) calculated with Satterthwaite approximation, num = numerator, den = denominator

As a result, all main effects, two-way and three-way interaction were significant (all F > 11.74, $p < .001$, see Table 3). Overall a similar pattern than for the reading aloud task was found: the main effect of Age was driven by the slow responses measured in 5th graders (as shown by the follow-up analyses), the main effect of Language indicated slower responses in French than in German, and the effect of Number Size revealed slower responses for large (*i.e.* \`70s, \`80s and \`90s) than small numbers (*i.e.* \`30s, \`40s and \`50s) All two-way interactions were significant (see Table 3) and we found a three-way interaction between Age, Language and Number Size.

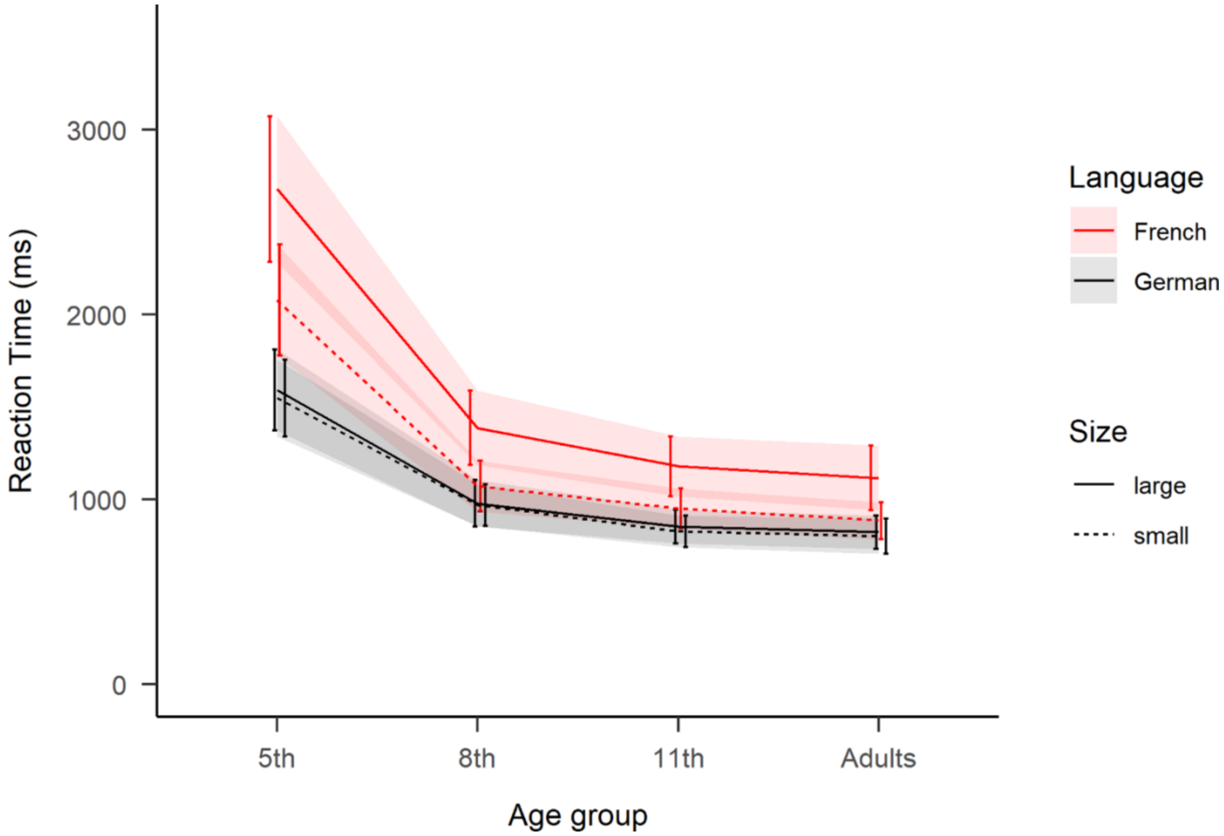


Fig 5. Reaction times in the verbal-visual matching task. For each age group as a function of "Language" and "Number Size". Large numbers correspond to '70s, '80s and '90s, while small numbers correspond to '30s, '40s and '50s. Ribbons represent standard error.

Follow-up analyses were applied comparing the estimated marginal means with paired comparisons on Satterthwaite corrected degrees of freedom and Bonferroni adjusted p-values. Follow-up analyses confirm the hypothesis relating to *transparency of power,* given a cost in French for '70s, '80s and '90s compared to '30s, '40s and '50s numbers for all age groups: 5th graders with a 682 ms cost ($t(97.71) = 10.84, p < .001$), 8th graders 339 ms cost ($t(53.31) = 5.27, p < .001$), 11th graders a 202 ms cost ($t(31.66) = 3.35, p < .01$), and adults with a 218 ms cost ($t(77.60) = 3.30, p < .01$). In contrast, the same comparison in German did not result in any significant differences ($all\ t < 1.05, n.s.$, max difference = 46 ms). In sum, the cost for vigesimal numbers observed in the reading aloud task is replicated with the verbal-visual matching task for all age groups, see Fig 5.

However, in comparison to the reading aloud task, the hypothesis on *language of math acquisition* tested on small numbers (*i.e.* `30s, `40s and `50s; see B in Fig. 1) was less supported, despite all differences in all four age-groups being positive. Indeed, performance advantages for German (LM1) were observed only in the youngest age group: while 5th graders were 682 ms slower in French than in German ($t(121.9) = 5.55, p < .001$), the same comparison between older age groups was not significant, 340 ms for 8th graders ($t(106.54) = 1.25, p = 1.0$), 202 ms for 11th graders ($t(107.1) = 1.34, p = 1.0$) and 218 ms for adults ($t(104.8) = 0.82, p = 1.0$). As for the previous reading aloud task, we compared with z-scores transformed and on a subsampled dataset to test the robustness of these results, see Fig 6.

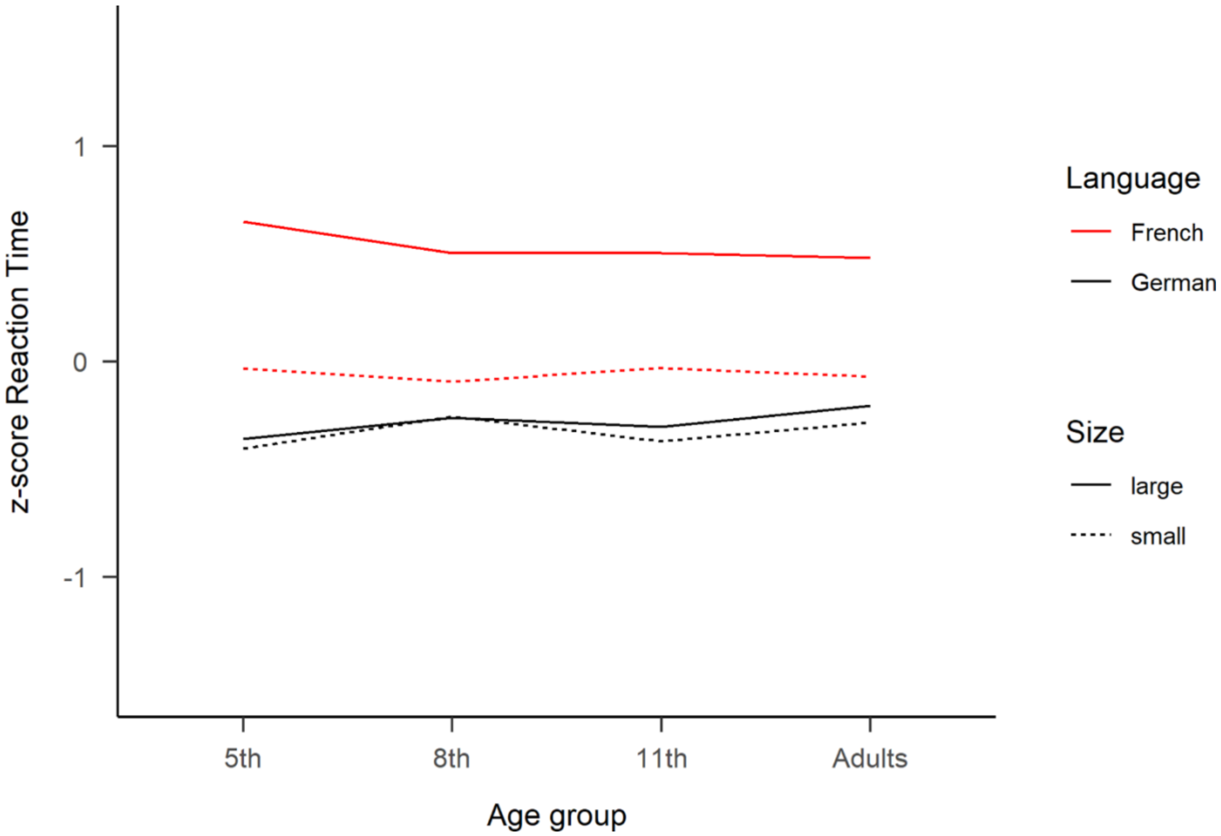


Fig 6. Z-score reaction times of the verbal-visual matching task. Standardized reaction times for each age group as a function of "Language" and "number Size". Large numbers correspond to '70s, '80s and '90s, while small numbers correspond to '30s, '40s and '50s.

The linear model applied on z-scores (see Fig 6) replicated the results with raw data described above, except, the three-way interaction with age, suggesting that the age differences observed with raw RTs, might be caused by differences between age-group variability (see S2 Table 3). The analyses on the subset with four syllable-long number words replicated the effects mainly for the younger age groups, as the number size effect revealed that a cost associated with vigesimal numbers was present for 5th and 8th graders. Moreover, the effect of language in favour of the first language of math acquisition (LM1, *i.e.* German) was significant only in 5th graders (See S2 Table 3).

4.3.2 Correct Responses

For the matching task, the correct responses were analysed with the same method and model as for the reading aloud task, namely a binomial approach and likelihood ratio test, see formula (B). Thus for the verbal-visual matching task, we modelled again the main effects and interactions between Language and Number Size and added random intercept per subject to consider individual differences.

Table 4. Results of the verbal-visual matching task's CR's linear mixed model

	df	χ^2	*p-value*
Language	1	22.31	<0.001
Number Size	1	3.20	0.074
Language x Number Size	1	16.23	<0.001

Similarly to the reading aloud task, the correct response rates showed a main effect of Language indicating in average 5.45 % more errors were done in French than in German, see Table 4. The main effect of Number Size was marginally significant, potentially indicating more errors for large numbers. The two-way interaction between Language and Number Size was also significant again, see Fig 7.

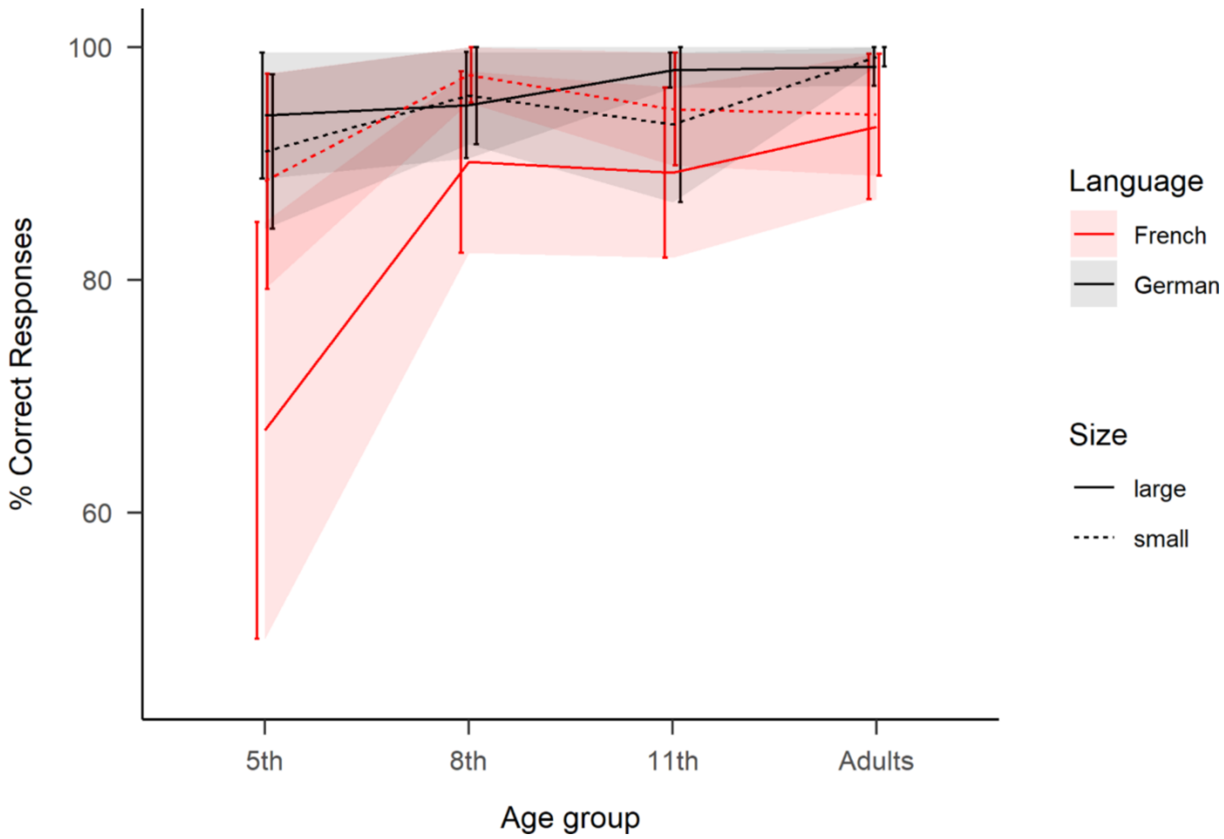


Fig 7. Correct Responses in the reading aloud task. Percent correct for each age group as a function of "Language" and "Number Size". Large numbers correspond to '70s, '80s and '90s, while small numbers correspond to '30s, '40s and '50s. Ribbons represent standard error.

Follow-up analyses indicate that 8.5 % more errors were made when matching the French vigesimal '70s, '80s and '90s numbers than '30s, '40s and '50s numbers ($z = -4.75, p < .001$), as would be expected from the effect of transparency of power. No such difference was found in German ($z = 1.37, p = 1.0$). Secondly, comparing small numbers in both languages revealed no differences between French (LM2) and German (LM1) ($z = -0.46, p = 1.0$), thus again failing to reveal effects of language of math acquisition in the verbal-visual matching task.

In summary, for the verbal-visual matching task the (1) *transparency of power* hypothesis could be confirmed, as transcoding performances were overall lower for French '70s, '80s and '90s numbers having a vigesimal structure, stably across age groups. Concerning the hypothesis on (2) *language of math acquisition,* the advantage for transcoding in German (LM1) was only stable across age groups, when considering standardized RTs (see Fig. 6).

5 Discussion

Transcoding speed and accuracy of two-digit numbers were measured in bilingual participants during a reading numbers aloud and verbal-visual matching task, with a transversal developmental design. Participants were German-French bilinguals from four age groups consisting of 5^{th}, 8^{th}, 11^{th} grades and adults. Since, the language of learning mathematics switches from German (LM1) to French (LM2) in 7^{th} grade in the Luxembourg education system, participants from older age groups were increasingly trained with numbers and math in their LM2. For the reading aloud task, numbers were transcoded from visual to verbal formats. In the verbal-visual matching task numbers were transcoded from verbal to visual formats, with an additional target selection among three distractors. The main strength of the present study was the within-subject design which permits to examine language effects, while limiting external influences such as inter-individual variability or differences in culture or education, typically present in cross-linguistic studies (Krinzinger et al., 2011). The use of the same task across all age groups allowed direct comparison of transcoding across different stages of bilingualism.

As expected, performance improved with age and education, such that participants from older age groups were faster and more accurate in transcoding two-digit numbers than younger one. Critically, independently of age and task, within French, there was a response time cost for transcoding `70s, `80s and `90s compared to `30s, `40s and `50s numbers. These results are in line with the concept of *transparency of power* since `70s, `80s and `90s number words in French follow a less transparent and hence more costly base-20 system than numbers under 60, which are characterized by a more transparent base-10 structure. This pattern was confirmed by analysing correct responses. However, age could not be included in the linear mixed model for accuracy, probably due to ceiling effects in the older age groups. In addition, and equally important, the results indicated a relative slow-down for LM2 (French) compared to LM1 (German) for numbers under 60 across ages in the reading aloud task. Correct responses were also affected by LM2 in the reading task, but for the same reasons as mentioned above we could not investigate age differences in accuracy. Since these effects were not consistently observed in the verbal-visual matching task, the hypothesis relating to *language of math acquisition* was mostly confirmed by the reading aloud task.

In sum, the results reveal two different linguistic effects on number transcoding: a cost for number words with weaker *transparency of power* (*i.e.* French vigesimal) and a cost for the later *language of mathematical learning* (*i.e.* the language which was not the language of math acquisition). In the following, we will discuss these effects sequentially.

5.1 Transparency of power

In French, \`70s, \`80s and \`90s numbers differ in their *transparency of power* (Bahnmueller et al., 2018) from \`30s, \`40s, \`50s numbers, with the former being in base-20. Comparing these numbers across both tasks, we found a cost for French \`70s, \`80s and \`90s compared to \`30s, \`40s and \`50s numbers across age groups and task. We interpret this as an effect of *transparency of power* that is independent of participant's age and bilingual proficiency level. An advantage for larger compared to smaller numbers could also be explained by a number size effect (Brysbaert, 1995). However, this interpretation can be excluded, because the same comparison in German did not reveal any differences which, under the untested assumption that 5th graders do not have mature magnitude representations, would speak against the prediction of semantic models. The effect was not explained by word length either (see (N. C. Ellis & Hennelly, 1980)), since it was replicated with all age groups except adults when analysing only number words having the same length of four syllables. Comparable results were observed in terms of correct responses, although the different age groups could not be compared due to insufficient model fit. Since the number of correct responses displayed ceiling effects for older participants in both tasks and languages, correct responses might lack sensitivity and is therefore not an ideal dependent variable for measuring language differences in the present study (Zuber et al., 2009).

The cost for less transparent number words structures is in line with previous results. Investigating participants from the same German-French bilingual population than the present study, Van Rinsveld and colleagues reported that arithmetical problems with numbers over 70 in French led to slower, less accurate solutions than in German both in visual and auditory modes (Van Rinsveld et al., 2015). Furthermore, our results replicate and extend previous transcoding studies, finding more errors in France-French (less transparent) than Wallonia-French (more transparent) speaking seven years old children (Seron & Fayol, 1994). Costs for base-20 numbers were also found for number reading and matching tasks when comparing 5th graders speaking French or English (*i.e.* base-10 numbers) (Van Rinsveld & Schiltz, 2016). The present study extends these findings by comparing different age groups of increasingly proficient German-French bilinguals performing the same transcoding tasks. Notably, the cost was found for both reading aloud and number matching tasks that rely on transcoding from visual to verbal processes and *vice-versa*. The results from the z-scores standardization suggest that the cost explained by a difference in *transparency of power* was generally and stably persisted from 5th grade up to adulthood. In other words the cost maintains across age groups, tasks and the associated degree of language proficiency. In the following, we give two possible accounts for the cost for transcoding French base-20 numbers.

From the perspective of cognitive models presented in the introduction, the cost for *transparency of power* in French would fit with Power and Dal Martello's semantic model if rules reflecting the vigesimal form are added, since more rules would mean slower production (Power & Dal Martello, 1990). Taking into account development, the asemantic ADAPT model (A Developmental Asemantic and Procedural model for Transcoding (Barrouillet et al., 2004)), would partly fit with our results. ADAPT was proposed for number dictation (*i.e.* verbal to visual Arabic transcoding), hence similar processes could have taken place in the verbal to visual matching and reading aloud tasks (*i.e.* visual Arabic to verbal) of the present study design. In ADAPT, French `70s, `80s and `90s require more rules and therefore also more processing time. However, the model further proposes a lexicalization through repetitions over development, *i.e.* leading to faster transcoding of `70s, `80s and `90s. Yet, when standardizing age groups variability through z-score transformations, we noted that numbers above 70 were similarly slower to transcode in all age groups, arguing against a lexicalization with training. Also, in answer to an interesting comment in the review process, we conducted additional analyses on reaction times (see S4), by adding all decades as levels for the Number Size factor (hence 8 levels from '30s to '90s). Those additional analyses confirm that for both tasks in French, reaction times become slower from the '70s decade onwards. The non-lexicalization of `70s, `80s and `90s vigesimal numbers in French might interact with their late formal acquisition (*i.e.* LM2 acquired at 7th grade, around 11-12 years old). Finally, the cost for `70s, `80s and `90s numbers also fits with Dotan and Freidman's recent transcoding model which suggests that reading `70s and `90s French number words requires additional irregular rules. When reading 75 for example, a number word frame would be structured by a decade and a teen instead of ten, then the ten frame filling would be changed from "7" to "6" (the filled frame would be: [6:tens] [15:teens] to give "*soixante-quinze*", literally "*six*ty-*fifteen*") (Dotan & Friedmann, 2018).

Speculatively, it could be argued that for French morpho-syntactically complex number words that are not lexicalized, multiple additional morphemes need to be retrieved (see *i.e.*

(Meeuwissen et al., 2003)). This could be the origin of the cognitive slow-down as it requires to retrieve more lexical morphemes (*i.e.* four for 97: “quatre”, “vingt”, “dix”, “sept”). Eventually, these morphemes, which are made of other numbers might interfere among each other and slow down RTs (*i.e.* proactive interference (Campbell, 1995; De Visscher & Noël, 2014)).

Please note that all three models would also generally predict slower responses for German number words above 12 due to the additional inversion rule in German. Such effects of *transparency of order* have been found in previous studies (Poncin et al., 2019; Steiner, Banfi, et al., 2021; Zuber et al., 2009), but it is possible these were masked in the present study by the effect of language of math acquisition described in the following section.

To sum up, the effect of *transparency of power* was confirmed by costs for French `70s, `80s and `90s vigesimal numbers, independently from age or task. The origin of this cost could be explained by the non-automatized, hence not lexicalized, transcoding process for these non-transparent numbers in French. However, since French is the second language of mathematical acquisition (LM2) of the current sample, it remains an empirical question whether this interpretation can be generalized to early learners of French.

5.2 Language of Mathematical Acquisition: LM1 vs. LM2

To assess whether and how *language of math acquisition* impacted transcoding in the different age groups, we compared performances in German (LM1) and French (LM2). To avoid the potentially confounding effect of *transparency of power* described above, only `30s, `40s and `50s in both languages were compared. We can also exclude that differences *in transparency of order* between German and French explain LM2 costs since this would have meant the opposite effect, that is slower responses in German (LM1, having an inverted number word structure, *e.g.* (Moeller, Zuber, et al., 2015; Pixner et al., 2011; Poncin et al., 2019; Steiner, Banfi, et al., 2021)), rather than in French (LM2, having a non-inverted number word structure). In line with our expectations, we observed costs during the reading aloud task in LM2 (French) compared to the LM1 (German); these costs were observed in the two youngest age groups in the analyses on RT, but they appeared at all ages when considering standardized response speed and four-syllable words. In the verbal-visual matching task, costs were consistently visible only in the analyses of standardized response times.

Our findings are in line with studies reporting qualitatively different arithmetic performance with LM+ compared to LM- in bilingual Filipino-English and Spanish-Basque participants, even if LM- corresponds to participants' mother language (Bernardo, 2001; Salillas & Wicha, 2012). Finally, they also match and extend the finding that solving addition problems was slower in LM2 than LM1 in participants coming from the same education system than in the present study (Van Rinsveld et al., 2015).

Interestingly, while the language of math acquisition impacted the number reading task, this was considerably less the case in the verbal-visual matching task. Lexically, the different pattern of results for the LM2 could be explained by different memory retrieval mechanisms (Vander Beken & Brysbaert, 2018) since the number reading task can be considered as a form of free recall while the matching task is more similar to a familiarity judgment. During free recall all possible number words can interfere with the retrieval of the correct number word, entailing a kind of lexical competition among the different verbal codes causing a cost for the less dominant language (LM2). In contrast, visual familiarity of the target number might underlie participants' responses during matching, weakening the role of language code(s) activation during this task. Similar task differences were indeed reported by Vander Beken and Brysbaert (Vander Beken & Brysbaert, 2018) in a study investigating the recall of text in university students' first and second languages. An additional explanation might rely on the nature of the stimuli. In the matching task, the number word input is already language-specific and therefore might be less susceptible to between-language lexical competition. While for the number reading task the identical visual Arabic digits might lead to a lexical competition between the LM1 and LM2.

To interpret the LM2 cost different theoretical perspectives can be taken. A possible interpretation can be derived from the ADAPT model of number transcoding (Barrouillet et al., 2004). LM1 (German) could be lexicalized (i.e. directly retrieved from long-term memory), while the LM2 (French), could rely on slower procedural rules, even for numbers under 60. In line with this view, weaker fMRI temporal lobe activation was observed when solving simple additions in LM2, proposedly reflecting less verbal retrieval than for LM1 additions (Van Rinsveld et al., 2017). Furthermore, since in ADAPT algorithmic rules are enacted by the short-term memory, it could potentially impact its capacity by using more resources (A. Baddeley, 2003; Camos, 2008) and in turn explain parts of the LM2 costs observed in the same bilingual population for exact arithmetic (Van Rinsveld et al., 2015). It is worthwhile noticing that this disadvantage for storing and accessing LM2 numbers in verbal short-term memory might also impact the results obtained in neuro-developmental diagnostic tests. Indeed, using LM2 numbers for tests such as the number span test of the Wechsler intelligence scale or for different tasks in dyscalculia batteries might hamper performance and lead to an underestimation of children's cognitive abilities.

The LM2 cost is also explainable from a psycho-linguistic perspective: here we present a syntactic and a lexical interpretation. Syntactically, the LM2 cost might result from an over-generalization of the LM1 syntactic inversion rule (see for example (Zuber et al., 2009)): transcoding number words in French would require inhibition of the inversion rule (over)learned from German. A lexical explanations can be found in more general bilingual models such as the bilingual interactive activation model (BIA+), predicting a response competition between both languages, with faster activation for the more frequently used language (Dijkstra & van Heuven, 2002). In this framework, the slower lexical retrieval of LM2 would also have detrimental impacts on short-term memory (N. C. Ellis & Hennelly, 1980; Gathercole & Baddeley, 1993) and hence arithmetic (Friso-van den Bos et al., 2013).

In summary, the performance was overall better for LM1 (German) compared to LM2 (French) `30s, `40s, `50s numbers, confirming the importance of language of math acquisition for two-digit number transcoding in bilinguals. These effects were stronger in the number reading task than in the verbal-visual matching task and could be interpreted from bilingual lexical or syntactic perspectives.

A limitation of the present study is that it did not allow to disentangle the role played by the language of math acquisition from the familiarity with number words in the two languages. Since the language used for early learning (LM1) is probably also the one used more frequently, both factors might be confounded. Yet, both processes are supposed to rely on different neuronal substrates (Fiebach et al., 2003). It is, however, noteworthy that math education extended for one year longer in LM2 (7 years of secondary education) than in LM1 (6 years of primary education), which might help to balance the frequencies in both languages.

Concerning the effect of transparency of power which we observed with `70s, `80s and `90s French numbers, we cannot exclude that part of the effect comes from the mixed nature of French number words. Since French contains both base-10 and base-20 numbers, it remains indeed to be determined whether similar effects would be observed with a language containing exclusively base-20 numbers (see for example (Salillas et al., 2015)). Likewise, further interpretation of the LM2 costs would benefit from a study assessing the effects of LM1 on transcoding with a language having simpler number word structures (e.g. English or Asian languages). Finally, it is still debated whether transcoding requires access to semantics or not (see (Barrouillet et al., 2004)). Further studies should investigate LM2 costs during number processing tasks systematically activating number semantics (e.g. in masked priming designs as in (Reynvoet et al., 2002)).

Taken together the present results confirm and extend the interactions between language and number processing observed in more complex tasks such as arithmetic problem solving, corroborating the role of language in numerical cognition.

6 Conclusion

The *transparency of power* consistently affected transcoding in bilinguals from the four age groups, ranging from grade 5 to adulthood. Base-20 number words in French were transcoded slower than base-10 words and this effect could not be explained by a semantic number size effect or the length of number words. Furthermore, *language of math acquisition* affected the speed with which Arabic numbers were named, such that transcoding in LM2 entailed a cost across age groups. This allows us to conclude that linguistic factors influence basic numerical tasks such as transcoding until adulthood.

6.1 Acknowledgments

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We are grateful to the staff in schools for opening their doors to our team. We also thank the subjects of all age groups for their participation

6.2 Author Contributions:

Conceived and designed the experiments: AVR, CS. Performed the experiments: AVR, AP. Analysed the data: RL. Wrote the paper: RL, CS.

7 Supporting information

7.1 S1 Table 1 to 4.

Mean Reaction Times and Correct Responses of both tasks across the factors Age group, Language and Size.

7.2 S1. Reading aloud task

7.2.1 S1. Reaction Times (in ms)

S1 Table 1: Reading aloud task, reaction times (ms)					
Language	Decades	Age			
		5th grade	8th grade	11th grade	Adults
French	‘30s, ‘40s, ‘50s	1233.39(55)	833.16(18)	817.43(19)	855.61(35)
	‘70s, ‘80s, ‘90s	1516.37(81)	1053.10(30)	1006.2(23)	1092.40(34)
German	‘30s, ‘40s, ‘50s	722.48(18)	655.56(13)	655.83(13)	697.26(20)
	‘70s, ‘80s, ‘90s	732.18(17)	659.61(11)	665.52(14)	688.69(15)
Total mean		978.67 (25)	788.65 (11)	777.93 (10)	822.59 (15)
Note: standard errors in parenthesis					

7.2.2 S1. Correct Responses (in %)

S1 Table 2: Reading aloud task, correct responses (%)					
Language	Decades	Age			
		5th grade	8th grade	11th grade	Adults
French	‘30s, ‘40s, ‘50s	91(3)	99(1)	98(1)	97(2)
	‘70s, ‘80s, ‘90s	78(4)	95(2)	98(1)	95(2)
German	‘30s, ‘40s, ‘50s	99(1)	99(1)	99(1)	99(1)
	‘70s, ‘80s, ‘90s	98(1)	100(0)	99(1)	99(1)
Total mean		92(1)	98(1)	98(1)	98(1)
Note: standard errors in parenthesis					

7.3 S1. Verbal-Visual matching task

7.3.1 S1. Reaction Times (in ms)

S1 Table 3: Verbal-visual matching task, reaction times (ms)					
Language	Decades	Age			
		5th grade	8th grade	11th grade	Adults
French	‘30s, ‘40s, ‘50s	1924.54(93)	1066.44(34)	953.01(28)	879.52(32)
	‘70s, ‘80s, ‘90s	2606.68(135)	1406.40(66)	1155.80(40)	1098.50(51)
German	‘30s, ‘40s, ‘50s	1554.04(51)	974.10(33)	824.88(23)	795.32(30)
	‘70s, ‘80s, ‘90s	1600.93(56)	971.09(37)	850.20(22)	826.84(31)
Total mean		1830.83 (42)	1095.63(23)	938.07 (15)	891.15 (19)
Note: standard errors in parenthesis					

7.3.2 S1. Correct Responses (in %)

S1 Table 4: Verbal-visual matching task, correct responses (%)					
Language	Decades	Age			
		5th grade	8th grade	11th grade	Adults
French	‘30s, ‘40s, ‘50s	89(2)	98(1)	95(2)	94(2)
	‘70s, ‘80s, ‘90s	69(4)	90(2)	89(3)	93(3)
German	‘30s, ‘40s, ‘50s	91(2)	96(2)	93(2)	99(1)
	‘70s, ‘80s, ‘90s	94(2)	95(2)	98(2)	98(1)
Total mean		87(1)	94(1)	94(1)	96(1)
Note: standard errors in parenthesis					

7.4 S2 Table 1 to 6.

Linear Mixed models on z-score transformed data and on four syllables long subsampled dataset of both tasks.

7.4.1 S2. Reading aloud task

7.4.1.1 S2. RT z-score

The same linear mixed model as for RTs was applied on *z-score* transformed RTs, applied separately for each age group. This transformation aimed to reduce the variability of RT between the four age groups. Note that the same dataset was used, namely after removing post error slipping.

As expected, the transformation disrupted the main effect of Age as well as the two-way interactions and three-way interactions involving Age. Otherwise, these analyses replicated the main effects of Language, Number Size and importantly the two-way interaction between Language and Number Size remained significant (see S1 Table 1).

S1 Table 1. Results of the reading aloud task's z-score transformed RT's linear model

	num *df*	den *df*	F	Pr(>F)
Age	3	94.91	0.03	0.99
Language	1	79.43	175.18	<0.001
Number Size	1	26.87	31.54	<0.001
Age x Language	3	92.07	0.68	0.56
Age x Number Size	3	100.95	0.57	0.64
Language x Number Size	1	21.95	52.45	<0.001
Age x Language x Number Size	3	1800.99	1.13	0.33

Note: num = numerator, den = denominator

Follow-up analysis of the interaction between Language and Number Size confirmed both hypotheses. In French the vigesimal '70s, '80s, and '90s numbers were named slower ($t(24.7) = 9.35, p < .001$) than '30s, '40s, and '50s numbers, in line with the expected effect of *transparency of power*. While the same comparison was not significant in German ($t(24.1) = 0.05, n.s.$).

Secondly, comparing '30s, '40s, and '50s numbers in both languages, revealed a cost for naming numbers in French (LM2) compared to German (LM1) ($t(55.4) = 7.40, p < .001$) (see S1 Table 2), as expected according to the hypothesis on the effect of *language of math acquisition*.

S1 Table 2. Reading aloud task's average z-scores

Language	Size	Age			
		5th grade	8th grade	11th grade	Adults
French	'30s, '40s, '50s	0.21	0.13	0.15	0.08
French	'70s, '80s, '90s	0.62	0.89	0.91	0.82
German	'30s, '40s, '50s	-0.53	-0.49	-0.50	-0.41
German	'70s, '80s, '90s	-0.52	-0.48	-0.46	-0.44

7.4.1.2 S2. Subset data: four-syllable length

A second control for the length of the number words consisted in removing all number words that were longer or shorter than four syllables. This led to a reduction of the dataset from 2080 (after the exclusion of incorrect responses and post-error trials) to 1661 measurement points. The same pattern of results as with the complete dataset was found, besides the three-way interaction between Age Language and Number Size that is here not significant anymore, see S1 Table 3.

S1 Table 3. Results of the reading aloud task's linear model on the subset data

	num *df*	den *df*	F	Pr(>F)
Age	3	92.68	14.76	<0.001
Language	1	64.02	100.80	<0.001
Number Size	1	17.78	21.69	<0.001
Age x Language	3	90.37	13.60	<0.001
Age x Number Size	3	118.80	0.63	0.56
Language x Number Size	1	13.03	26.02	< 0.001
Age x Language x Number Size	3	1388.17	1.07	0.36

Note: num = numerator, den = denominator

Follow-up for the two-way interaction between Age and Language indicated a significant improvement in French between 5th and 8th graders ($t(97.3) = 5.73, p < .001$, but no significant changes in German between the four age groups (all $t < 2.13, p < .21$).

Follow-up analyses on the two-way interaction between Language and Number Size confirmed the cost, in French ($t(11.7) = 5.47, p < .001$) for the vigesimal '70s, '80s, and '90s numbers. The same comparison was not significant in German ($t(20.8) = 0.10, p = n.s.$).

Secondly, comparing '30s, '40s, and '50s numbers in both languages, again confirmed the cost for naming numbers in French (LM2) compared to German (LM1) ($t(47.6) = 6.22, p < .001$).

7.4.2 S2. Verbal-Visual Matching task

7.4.2.1 S2. RT z-score

The same linear mixed model as for RTs was again applied on *z-score* transformed RTs, applied separately for each age group. Confirming the raw data analyses, the main effects of Language, Number Size, and their interactions remained significant, while disrupting the main effect of Age and the interactions with this factor (see S1 Table 4).

S1 Table 4. Results of the matching task's z-score transformed RT's linear model

	num *df*	den *df*	F	Pr(>F)
Age	3	94.26	0.12	0.95
Language	1	67.53	105.12	<.001
Number Size	1	24.63	24.59	<.001
Age x Language	3	91.60	3.31	0.03
Age x Number Size	3	93.00.	0.47	0.71
Language x Number Size	1	21.81	60.55	<.001
Age x Language x Number Size	3	1967347	0.34	0.80

Note: num = numerator, den = denominator

Follow-up contrast analyses also replicated the results obtained with raw RTs, namely a cost for'70s, '80s, and '90s numbers in French ($t(23.6) = 7.21, p < .001$), but not for the same numbers in German ($t(24.9) = 0.93, p = n.s.$). Secondly, the comparison between '30s, '40s, and '50s in both languages revealed a cost for French (LM2) compared to German (LM1) ($t(51.0) = 4.49, p < .001$), see S1 table 5.

S1 Table 5. Matching task's average z-scores

Language	Size	Age			
		5th grade	8th grade	11th grade	Adults
French	'30s, '40s, '50	-0.03	-0.09	-0.03	-0.07
French	'70s, '80s, '90	0.65	0.50	0.50	0.48
German	'30s, '40s, '50	-0.41	-0.25	-0.37	-0.28
German	'70s, '80s, '90	-0.36	-0.26	-0.30	-0.20

7.4.2.2 S2. Subset data: four-syllable length

The removal of items corresponding to number words with more or less than 4 syllables reduced the sample from 1998 to 1499 measurement points. Like with the complete dataset, the model replicated all significant main effects and interactions (all F > 10.42, p < .001)(S1 Table 6).

S1 Table 6. Results of the matching task's linear model on the subset data

	num *df*	den *df*	F	Pr(>F)
Age	3	93.66	53.02	<.001
Language	1	74.64	56.78	<.001
Number Size	1	22.47	17.75	<.001
Age x Language	3	91.58	11.32	<.001
Age x Number Size	3	89.84	7.69	<.001
Language x Number Size	1	16.67	27.38	<.001
Age x Language x Number Size	3	1399.65.	7.07	<.001

Note: num = numerator, den = denominator

Follow-up analyses on the three-way interaction between Age, Language, and Number Size revealed slightly different results than those obtained with the complete data set, in particular for the two older age. Indeed, the follow-up contrasts in French resulted in a cost for vigesimal ’70s, ’80s, and ’90s numbers compared to ’30s, ’40s, and ’50s in 5th graders ($t(48.19) = 8.33, p < .001$),8th graders ($t(40.19) = 3.92, p < .01$). and 11th graders ($t(4488) = 2.77, p = .05$). This contrast was however not significant in adults ($t(58.29) = 2.24, p = .17$). Furthermore, as in previous analyses, the same comparison in German, did not lead to significant differences (all $t < .59, p = n.s.$). Our hypothesis on the cost entailed by transcoding vigesimal numbers (i.e. French ’70s, ’80s, and ’90s) was thus only confirmed for the two youngest age groups with this reduced data set.

Secondly, when comparing ’30s, ’40s, and ’50s numbers in French (LM1) and German (LM1), only 5th graders showed the expected cost in French ($t(125.83) = 5.25, p < .001$). This was not the case for 8th and 11th graders and adults (all $t < 1.30, p = n.s.$). Analyses on four-syllables number words did not fully support the hypothesis on the language of math acquisition. In conclusion, with the reduced four-syllable data set we observed the effect of number word transparency was robustly in 5th and 8th graders, while the LM2 cost was robust only in 5th graders.

7.5 S3 Table 1: Stimuli.

Stimuli and number of syllables used in the experiment. Number of syllables for each number word in German and in French as used to control for word length.

Digit	German	Syllables	Phono L	French	Syllables	Phono L	Task
34	vierunddreissig	4	11	trentequatre	4	8	Read
35	fünfunddreissig	4	12	trentecinq	3	7	Read
36	sechsunddreissig	4	12	trentesix	3	7	Read
38	achtunddreissig	4	11	trentehuit	3	7	Read
41	einundvierzig	4	11	quaranteetun	4	7	Read
43	dreiundvierzig	4	12	quarantetrois	4	9	Read
45	fünfundvierzig	4	13	quarantecinq	4	8	Read
46	sechsundvierzig	4	13	quarantesix	4	8	Read
51	einundfünfzig	4	12	cinquanteetun	4	7	Read
52	zweiundfünfzig	4	13	cinquantedeux	4	8	Read
57	siebenundfünfzig	5	15	cinquantesept	4	8	Read
59	neunundfünfzig	4	13	cinquanteneuf	4	8	Read
61	einundsechzig	4	11	soixanteetun	4	8	Read
63	dreiundsechzig	4	12	soixantetrois	4	10	Read
65	fünfundsechzig	4	13	soixantecinq	4	9	Read
69	neunundsechzig	4	12	soixanteneuf	4	9	Read
72	zweiundsiebzig	4	12	soixantedouze	5	9	Read
74	vierundsiebzig	4	12	soixantequatorze	5	13	Read
78	achtundsiebizig	4	12	soixantedixhuit	5	12	Read
79	neunundsiebzig	4	12	soixantedixneuf	5	12	Read
81	einundachtzig	4	10	quatrevingtun	4	7	Read
83	dreiundachtzig	4	11	quatrevingttrois	4	10	Read
84	vierundachtzig	4	11	quatrevingtquatre	5	10	Read
86	sechsundachtzig	4	12	quatrevingtsix	5	9	Read
93	dreiundneunzig	4	12	quatrevingttreize	4	10	Read
95	fünfundneunzig	4	13	quatrevingtquinze	4	9	Read
96	sechsundneunzig	4	13	quatrevingtseize	4	9	Read
98	achtundneunzig	4	12	quatrevingtdixhuit	5	12	Read
31	einunddreissig	4	10	trenteetun	3	6	Match
32	zweiunddreissig	4	11	trentedeux	3	6	Match
37	seiebenunddreissig	5	13	trentesept	3	7	Match
39	neununddreissig	4	11	trenteneuf	3	7	Match
42	zweiundvierzig	4	12	quarantedeux	4	7	Match
47	siebenundvierzig	5	14	quarantesept	4	8	Match
48	achtundvierzig	4	12	quarantehuit	4	8	Match
49	neunundvierzig	4	12	quaranteneuf	4	8	Match
53	dreiundfünfzig	4	13	cinquanetrois	4	9	Match
54	vierundfünfzig	4	13	cinquantequatre	5	9	Match

56	sechsundfünfzig	5	14	cinquantesix	4	8	Match
58	achtundfünfzig	4	13	cinquantehuit	4	8	Match
62	zweiundsechzig	4	12	soixantedeux	4	8	Match
64	vierundsechzig	4	12	soixantequatre	5	10	Match
67	siebenundsechzig	5	14	soixantesept	4	9	Match
68	achtundsechzig	4	12	soixantehuit	4	9	Match
71	einundsiebzig	4	11	soixanteetonze	4	8	Match
73	dreiundsiebzig	4	12	soixantetreize	4	10	Match
75	fünfundsiebizig	4	13	soixantequinze	4	9	Match
76	sechsundsiebzig	4	13	soixanteseize	4	9	Match
82	zweiundachtzig	4	11	quatrevingtdeux	4	8	Match
85	fünfundachtzig	4	12	quatrevingtcinq	4	9	Match
87	siebenundachtzig	5	13	quatrevingtsept	4	9	Match
89	neunundachtzig	4	11	quatrevingtneuf	4	9	Match
91	einundneunzig	4	11	quatrevingtonze	4	8	Match
92	zweiundneunzig	4	12	quatrevingtdouze	4	9	Match
94	vierundneunzig	4	12	quatrevingtquatorz e	5	13	Match
97	siebenundneunzig	5	14	quatrevingtdixsept	5	12	Match

Note: Read = reading aloud task. Match = verbal-visual matching task. Phonological lengths (PhonoL) were retrieved from CLEARPOND (Marian et al., 2012)

7.6 S4. Supplementary Analyses per decades:

Analyses carried on each decades. Linear mixed models for both tasks' reaction times, but replacing the Number Size factor's level with decades from '30s until '90s (instead of small and large).

In the following additional analyses with linear mixed models, we compared all decades. That is, in the following, the factor *Number Size*, has **8 levels: '30s, '40s, '50s, '60s, '70s, '80s, '90s**. We used exactly the same model for reaction times as described in the results (i.e. (A)) on the data removing the post-error slow down. All degrees of freedom are calculated with Satterthwaite approximation.

7.6.1 S4. Reading aloud task:

S4 Table 1: results of the linear mixed model *per decades*

	num Df	den Df	F	Pr (>F)
Age	3	93.15	17.64	< 0.001
Language	1	97.84	682.33139.44	<0.001
Number Size	6	20.93	6.41	<0.001
Age x Language	3	90.09	17.48	<0.001
Age x Number Size	18	2155.72	2.84	0.005
Language x Number Size	6	21.00	13.35	< 0.001
Age x Language x Number Size	18	2155.15	2.04	0.006

Note: Number Size has 8 levels: '30s, '40s, '50s, '60s, '70s, '80s, '90s.

7.6.1.1 S4. "Stepwise" contrasts:

The following custom contrasts compare each decade with the previous and following one. For example '30 vs '40, '40 vs '50, etc. Contrasts are calculated on estimated marginal means, degrees of freedom from Satterthwaite approximation and p-values are Bonferroni corrected.

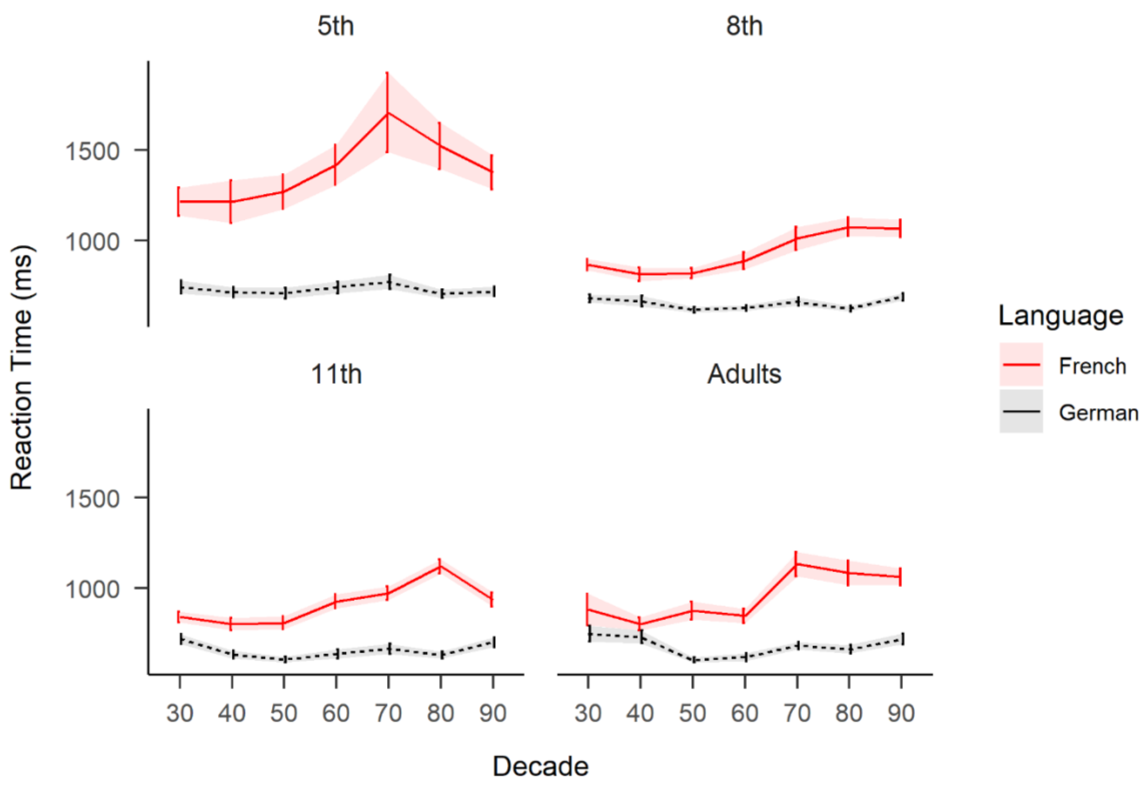
S4 Table 2: Contrasts between subsequent decades

French						
Age	contrast	estimate	SE	df	t.ratio	p.value
5th grade	30 - 40	-60.74	64.20	68.00	-0.95	0.35
	40 - 50	-75.19	64.30	68.60	-1.17	0.25
	50 - 60	-57.31	64.50	69.30	-0.89	0.38
	60 - 70	-386.08	70.70	98.30	-5.46	<.0001

	70 - 80	227.47	71.5 1	2.60	3.18	0.001
	80 - 90	-17.00	67.10	80.70	-0.25	0.80
8th grade	30 - 40	70.38	58.2	46.3	1.209	0.2328
	40 - 50	-17.61	58.1	45.9	-0.303	0.7632
	50 - 60	-48.96	58.1	46	-0.842	0.4042
	60 - 70	-147.22	60.1	52.5	-2.448	0.0177
	70 - 80	-50.7	60.9	55.3	-0.832	0.409
	80 - 90	17.12	59.4	50	0.288	0.7743
11th grade	30 - 40	42.62	59.9	51.7	0.712	0.4796
	40 - 50	-26.24	60	52.1	-0.437	0.6636
	50 - 60	-100.27	59.6	50.6	-1.684	0.0984
	60 - 70	-61.74	60.6	54.3	-1.018	0.313
	70 - 80	-114.42	63	63.1	-1.816	0.074
	80 - 90	168.64	61.6	57.8	2.738	0.0082
Adults	30 - 40	83.6	64.3	68.6	1.3	0.198
	40 - 50	-58.38	64.2	68.1	-0.909	0.3664
	50 - 60	15.68	64	67.1	0.245	0.807
	60 - 70	-290.96	65.6	74.1	-4.435	<.0001
	70 - 80	44.51	68.5	87.5	0.65	0.5175
	80 - 90	19.82	66.7	78.8	0.297	0.7671
German						
Age	contrast	estimate	SE	df	t.ratio	p.value
5th grade	30 - 40	27.98	56.9	56	0.492	0.6248
	40 - 50	5.04	56.9	56.1	0.089	0.9297
	50 - 60	-33.46	57.6	58.8	-0.581	0.5636
	60 - 70	-21.63	57.7	59.1	-0.375	0.709
	70 - 80	54.19	56.9	56.1	0.953	0.3449
	80 - 90	-13.34	56.7	55.5	-0.235	0.815
8th grade	30 - 40	15.67	55.5	50.8	0.282	0.7788
	40 - 50	47.99	55.3	50.1	0.868	0.3896
	50 - 60	-11.74	55.2	49.7	-0.213	0.8323
	60 - 70	-33.47	55.2	49.7	-0.607	0.5468
	70 - 80	40.89	55.7	51.7	0.734	0.4666
	80 - 90	-67.49	55.8	51.8	-1.211	0.2316
11th grade	30 - 40	71.35	57.6	58.9	1.238	0.2205
	40 - 50	38.87	58.1	61	0.669	0.506
	50 - 60	-34.06	57.4	58	-0.594	0.5549
	60 - 70	-21.73	57.4	58	-0.379	0.7062
	70 - 80	35.96	57.2	57.3	0.629	0.532
	80 - 90	-72.91	57.4	58.1	-1.271	0.2089

Adults	30 - 40	19.18	61.3	75.4	0.313	0.7554
	40 - 50	114.77	62	78.6	1.852	0.0678
	50 - 60	-10.61	62.5	81	-0.17	0.8656
	60 - 70	-54.77	62.3	80.1	-0.879	0.382
	70 - 80	17.14	61.2	74.6	0.28	0.78
	80 - 90	-56.22	60.9	73.3	-0.924	0.3587

In French, ‘30s, ‘40s, ‘50s numbers have comparable RTs. However, ‘60s have significant faster RT than ‘70s for 5th, 8th and adults. 11th graders show a marginally significant (p = .07) slow-down between ‘70s and ‘80s. None of those contrasts are significant in German.



S4 Fig1. Mean reaction time of the reading aloud task for each decade at each age groups. Ribbons represent one standard error.

7.6.2 S4. Verbal-visual matching

S4 Table 3: results of the linear mixed model per decades				
	num Df	den Df	F	Pr(>F)
Age	3	91.40	50.99	<0.001
Language	1	83.23	66.66	<0.001
Number Size	6	20.21	6.62	<0.001
Age x Language	3	85.79	15.35	<0.001
Age x Number Size	18	2017.63	3.62	<0.001
Language x Number Size	6	20.35	9.05	<0.001
Age x Language x Number Size	18	2017.28	2.75	<0.001

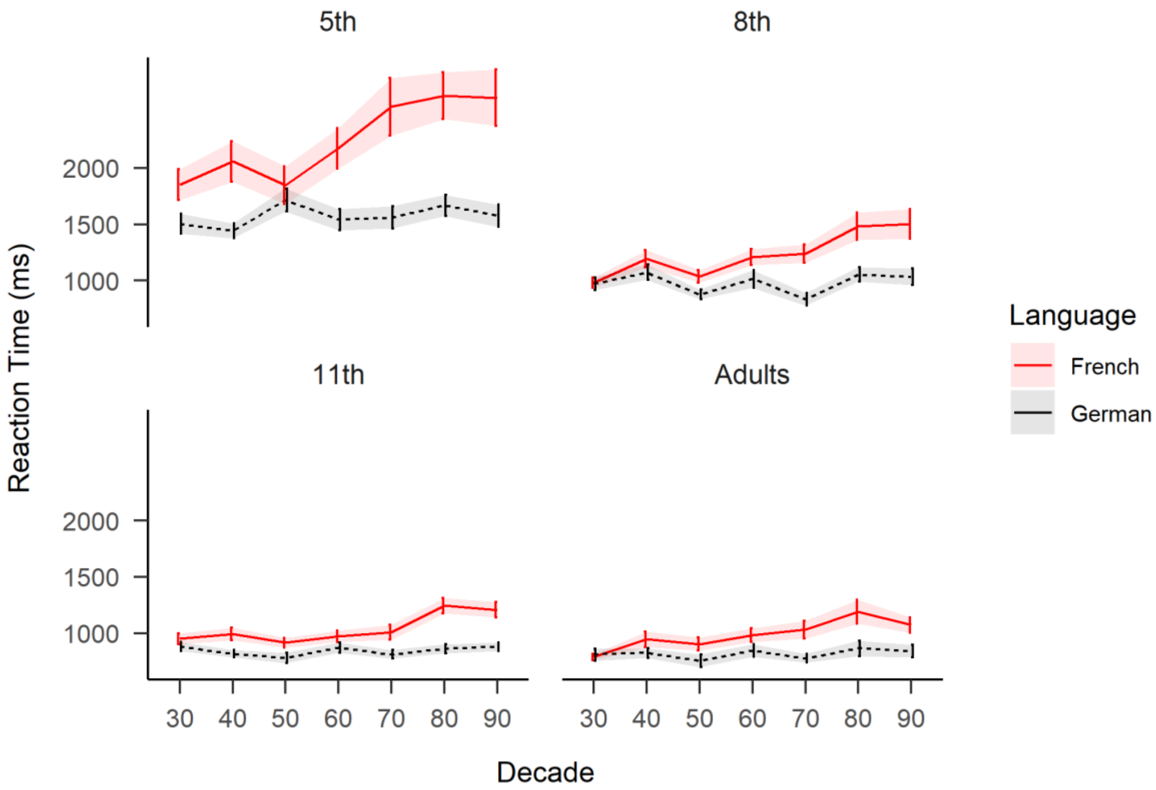
7.6.2.1 S4. "Stepwise" contrasts:

The following custom contrasts compare each decade with the previous and following one. For example '30 vs '40, '40 vs '50, etc. Contrasts are calculated on estimated marginal means, degrees of freedom from Satterthwaite approximation and p-values are Bonferroni corrected.

S4 Table 4: Contrasts between subsequent decades						
French						
	contrast	estimate	SE	df	t.ratio	p.value
5th grade	30 – 40	-183.46	117.50	80.60	-1.56	0.1223
	40 - 50	114.62	119.90	87.00	0.96	0.3416
	50 - 60	-435.53	123.70	97.50	-3.52	0.0007
	60 - 70	-381.76	130.40	119.40	-2.93	0.0041
	70 - 80	-22.69	134.50	134.10	-0.17	0.8663
	80 - 90	-17.12	131.60	124.10	-0.13	0.8967
8th grade	30 – 40	-173.54	105.6	52.8	-1.644	0.1062
	40 - 50	133.03	105.7	53.1	1.258	0.2137
	50 - 60	-117.65	105.5	52.8	-1.115	0.2698
	60 - 70	-121.56	107.7	57.2	-1.129	0.2636
	70 - 80	-183.97	107.9	57.8	-1.705	0.0936
	80 - 90	-32.91	108.3	58.4	-0.304	0.7623
11th grade	30 – 40	-18.24	107.6	57.1	-0.17	0.866
	40 - 50	52.74	107.3	56.5	0.491	0.6251
	50 - 60	-67.36	110.4	63	-0.61	0.5438
	60 - 70	-53.57	114.1	71.5	-0.47	0.64
	70 - 80	-205.86	112.6	68	-1.829	0.0718
	80 - 90	30.54	112.4	67.8	0.272	0.7868

Adults	30 – 40	-177.21	115.7	76	-1.532	0.1297
	40 - 50	61.05	115.6	75.7	0.528	0.5988
	50 - 60	-100.44	115.6	75.7	-0.869	0.3877
	60 - 70	-58.69	116.1	76.7	-0.506	0.6146
	70 - 80	-135.44	118.4	82.7	-1.144	0.256
	80 - 90	109.33	121.5	91.8	0.9	0.3705
German						
5th grade	contrast	estimate	SE	df	t.ratio	p.value
	30 – 40	63.01	103.4	76.8	0.61	0.5439
	40 - 50	-245.4	104.2	79.2	-2.355	0.021
	50 - 60	161.73	104.1	78.8	1.554	0.1242
	60 - 70	-50.87	102.7	74.7	-0.496	0.6216
	70 - 80	-83.95	103.1	75.9	-0.815	0.4178
	80 - 90	79.17	104	78.7	0.761	0.4489
8th grade	30 – 40	-115.7	99.3	65.5	-1.165	0.2483
	40 - 50	183.35	99	64.6	1.853	0.0685
	50 - 60	-119.82	99.4	65.7	-1.205	0.2325
	60 - 70	168.74	98.9	64.4	1.706	0.0928
	70 - 80	-202.7	98.2	62.7	-2.064	0.0432
	80 - 90	-10.12	99.5	66.1	-0.102	0.9193
11th grade	30 – 40	47.55	103.4	76.9	0.46	0.6469
	40 - 50	37.87	103.5	77.2	0.366	0.7155
	50 - 60	-91.99	103.7	77.8	-0.887	0.3776
	60 - 70	56.91	102.1	73.3	0.557	0.579
	70 - 80	-43.62	100.7	69.4	-0.433	0.6663
	80 - 90	-31.17	101.9	72.5	-0.306	0.7605
Adults	40	-33.83	106.9	87.6	-0.317	0.7523
	50	88.15	106.9	87.6	0.825	0.4116
	60	-88.4	106.9	87.6	-0.827	0.4103
	70	71.96	107.9	90.8	0.667	0.5064
	80	-81.46	108.8	94	-0.748	0.456
	90	-1.27	108.9	94.2	-0.012	0.9907

In French, ‘30s, ‘40s, ‘50s numbers have comparable RTs. However, ‘60s have significant faster RT than ‘70s for 5th graders and ‘70s are faster than ‘80s fro 8th, and 11th graders. No differences were found for adults. In German the only significant differences found were between ‘40s and ‘50s in 5th graders and between ‘70s and ’80s in 8th graders.



S4 Fig. 2: Mean reaction time of the verbal-visual matching task for each decade at each age groups. Ribbons represent one standard error.

8 References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*, *128*(3), 466–478. https://doi.org/10.1016/j.actpsy.2008.03.014

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Bahnmueller, J., Moeller, K., Mann, A., & Nuerk, H.-C. (2015). On the limits of language influences on numerical cognition – no inversion effects in three-digit number magnitude processing in adults. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01216

Bahnmueller, J., Nuerk, H.-C., & Moeller, K. (2018). A Taxonomy Proposal for Types of Interactions of Language and Place-Value Processing in Multi-Digit Numbers. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01024

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00328

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004a). ADAPT: A Developmental, Asemantic, and Procedural Model for Transcoding From Verbal to Arabic Numerals. *Psychological Review*, *111*(2), 368–394. https://doi.org/10.1037/0033-295X.111.2.368

Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004b). ADAPT: A Developmental, Asemantic, and Procedural Model for Transcoding From Verbal to Arabic Numerals. *Psychological Review*, *111*(2), 368–394. https://doi.org/10.1037/0033-295X.111.2.368

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*(3), 201–221. https://doi.org/10.1016/S0010-0277(02)00178-6

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bernardo, A. B. I. (2001). Asymmetric activation of number codes in bilinguals: Further evidence for the encoding complex model of number processing. *Memory & Cognition*, *29*(7), 968–976. https://doi.org/10.3758/BF03195759

Bernardo, A. B. I. (2005). Language and Modeling Word Problems in Mathematics Among Bilinguals. *The Journal of Psychology*, *139*(5), 413–425. https://doi.org/10.3200/JRLP.139.5.413-425

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General*, *124*(4), 434–452. https://doi.org/10.1037/0096-3445.124.4.434

Bürkner, P.-C. (2017). brms: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Bylund, E., Antfolk, J., Abrahamsson, N., Olstad, A. M. H., Norrman, G., & Lehtonen, M. (2023). Does bilingualism come with linguistic costs? A meta-analytic review of the bilingual lexical deficit. *Psychonomic Bulletin & Review*, *30*(3), 897–913. https://doi.org/10.3758/s13423-022-02136-7

Camos, V. (2008). Low working memory capacity impedes both efficiency and learning of number transcoding in children. *Journal of Experimental Child Psychology*, *99*(1), 37–57. https://doi.org/10.1016/j.jecp.2007.06.006

Campbell, J. (1995). Mechanisms of Simple Addition and Multiplication: A Modified Network-interference Theory and Simulation. *Mathematical Cognition*, *1*, 121–164.

Campbell, J. I. D., & Epp, L. J. (2004). An Encoding-Complex Approach to Numerical Cognition in Chinese-English Bilinguals. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *58*(4), 229–244. https://doi.org/10.1037/h0087447

Cerda, V. R., Grenier, A. E., & Wicha, N. Y. Y. (2019). Bilingual children access multiplication facts from semantic memory equivalently across languages: Evidence from the N400. *Brain and Language*, *198*, 104679. https://doi.org/10.1016/j.bandl.2019.104679

Chincotta, D., & Underwood, G. (1996). Mother tongue, language of schooling and bilingual digit span. *British Journal of Psychology*, *87*(2), 193–208. https://doi.org/10.1111/j.2044-8295.1996.tb02585.x

Chincotta, D., & Underwood, G. (1997). Speech Rate Estimates, Language of Schooling and Bilingual Digit Span. *European Journal of Cognitive Psychology*, *9*(3), 325–348. https://doi.org/10.1080/713752562

Chrisomalis, S. (2010). *Numerical Notation: A Comparative History*. Cambridge University Press.

Clayton, F. J., Copper, C., Steiner, A. F., Banfi, C., Finke, S., Landerl, K., & Göbel, S. M. (2020). Two-digit number writing and arithmetic in Year 1 children: Does number word inversion matter? *Cognitive Development*, *56*, 100967. https://doi.org/10.1016/j.cogdev.2020.100967

Cohen, J., Mac Whinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers*, *25*(2), 257–271.

Colomé, àngels, Laka, I., & Sebastián-Gallés, N. (2010). Language effects in addition: How you say it counts. *Quarterly Journal of Experimental Psychology*, *63*(5), 965–983. https://doi.org/10.1080/17470210903134377

Colomé, À. (2001). Lexical Activation in Bilinguals' Speech Production: Language-Specific or Language-Independent? *Journal of Memory and Language*, *45*(4), 721–736. https://doi.org/10.1006/jmla.2001.2793

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589–608. https://doi.org/10.1037/0033-295X.100.4.589

Comrie, B. (2013). Numeral bases (v2020.3). In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Zenodo. https://doi.org/10.5281/zenodo.7385533

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002

de Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. Psychology Press. https://doi.org/10.4324/9780203841228

De Visscher, A., & Noël, M.-P. (2014). The detrimental effect of interference in multiplication facts storing: Typical development and individual differences. *Journal of Experimental Psychology: General*, *143*(6), 2380–2400. https://doi.org/10.1037/xge0000029

De Vos, T. (1992). Tempo test rekenen (TTR)[Arithmetic number fact test]. *Nijmegen: Berkhout*.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*(1), 1–29. https://doi.org/10.1016/0010-0277(92)90030-l

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence. *Science*, *284*(5416), 970–974. https://doi.org/10.1126/science.284.5416.970

Del Maschio, N., & Abutalebi, J. (2019). Language Organization in the Bilingual and Multilingual Brain. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the Neuroscience of Multilingualism* (pp. 197–213). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119387725.ch9

Deloche, G., & Seron, X. (1982). From three to 3: A differential analysis of skills in transcoding quantities between patients with Broca's and Wernicke's aphasia. *Brain*, *105*(4), 719–733. https://doi.org/10.1093/brain/105.4.719

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, *82*(1), 64–81. https://doi.org/10.1348/2044-8279.002002

Dewaele, J.-M. (2007). Multilinguals' language choice for mental calculation. *Intercultural Pragmatics*, *4*(3). https://doi.org/10.1515/IP.2007.017

Dijkstra, T. (2005). Bilingual Visual Word Recognition and Lexical Access. In *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). Oxford University Press. https://doi.org/10.1017/S0272263107210071

Dijkstra, T., & Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Wahl, A., Buytenhuijs, F., Halem, N. V., Al-Jibouri, Z., Korte, M. D., & Rekké, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Dotan, D., & Friedmann, N. (2018). A cognitive model for multidigit number reading: Inferences from individuals with selective impairments. *Cortex*, *101*, 249–281. https://doi.org/10.1016/j.cortex.2017.10.025

Dowker, A., & Nuerk, H.-C. (2016). Editorial: Linguistic Influences on Mathematics. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01035

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Duyck, W., & Brysbaert, M. (2002). What number translation studies can teach us about the lexico-semantic organisation in bilinguals. *Psychologica Belgica*, *42*(3), 151–175.

Duyck, W., & Brysbaert, M. (2004). Forward and Backward Number Translation Requires Conceptual Mediation in Both Balanced and Unbalanced Bilinguals. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 889–906. https://doi.org/10.1037/0096-1523.30.5.889

Duyck, W., Depestel, I., Fias, W., & Reynvoet, B. (2008). Cross-lingual numerical distance priming with second-language number words in native- to third-language number word translation. *Quarterly Journal of Experimental Psychology*, *61*(9), 1281–1290. https://doi.org/10.1080/17470210802000679

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of Acquisition Effects in Adult Lexical Processing Reflect Loss of Plasticity in Maturing Systems: Insights From Connectionist Networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1103–1123. https://doi.org/10.1037/0278-7393.26.5.1103

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology*, *71*(1), 43–51. https://doi.org/10.1111/j.2044-8295.1980.tb02728.x

Fiebach, C. J., Friederici, A. D., Müller, K., von Cramon, D. Y., & Hernandez, A. E. (2003). Distinct brain representations for early and late learned words. *NeuroImage*, *19*(4), 1627–1637. https://doi.org/10.1016/S1053-8119(03)00227-1

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: A unified JavaScript framework for online studies. *International Conference on Computational Social Science (Cologne)*. https://www.labvanced.com/

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*(3), 819–824. https://doi.org/10.1016/j.cognition.2008.04.007

Frenck-Mestre, C., & Vaid, J. (1993). Activation of number facts in bilinguals. *Memory & Cognition*, *21*(6), 809–818. https://doi.org/10.3758/BF03202748

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/10.1016/j.edurev.2013.05.003

Garcia, O., Faghihi, N., Raola, A. R., & Vaid, J. (2021). Factors influencing bilinguals’ speed and accuracy of number judgments across languages: A meta-analytic review. *Journal of Memory and Language*, *118*, 104211. https://doi.org/10.1016/j.jml.2020.104211

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, *8*(3), 259. https://doi.org/10.1007/BF03174081

Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of Arithmetical Competencies in Chinese and American Children: Influence of Age, Language, and Schooling. *Child Development*, *67*(5), 2022–2044. https://doi.org/10.1111/j.1467-8624.1996.tb01841.x

Göbel, S. M., Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H.-C. (2014). Language affects symbolic arithmetic in children: The case of number word inversion. *Journal of Experimental Child Psychology*, *119*, 17–25. https://doi.org/10.1016/j.jecp.2013.10.001

Göbel, S. M., Watson, S. E., Lervåg, A., & Hulme, C. (2014). Children’s Arithmetic Development: It Is Number Knowledge, Not the Approximate Number Sense, That Counts. *Psychological Science*, *25*(3), 789–798. https://doi.org/10.1177/0956797613516471

Grabner, R. H., Saalbach, H., & Eckstein, D. (2012a). Language-Switching Costs in Bilingual Mathematics Learning. *Mind, Brain, and Education*, *6*(3), 147–155. https://doi.org/10.1111/j.1751-228X.2012.01150.x

Grabner, R. H., Saalbach, H., & Eckstein, D. (2012b). Language-Switching Costs in Bilingual Mathematics Learning. *Mind, Brain, and Education*, *6*(3), 147–155. https://doi.org/10.1111/j.1751-228X.2012.01150.x

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

Greisen, M., Georges, C., Hornung, C., Sonnleitner, P., & Schiltz, C. (2021). Learning mathematics with shackles: How lower reading comprehension in the language of mathematics instruction accounts for lower mathematics achievement in speakers of different home languages. *Acta Psychologica*, *221*, 103456. https://doi.org/10.1016/j.actpsy.2021.103456

Grosjean, F. (2001). *The Bilingual's Language Modes*. Blackwell Publishing.

Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.

Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.

Hahn, C. G. K., Saalbach, H., & Grabner, R. H. (2017). Language-dependent knowledge acquisition: Investigating bilingual arithmetic learning. *Bilingualism: Language and Cognition*, *22*(1), 47–57. https://doi.org/10.1017/S1366728917000530

Hahn, C. G. K., Saalbach, H., & Grabner, R. H. (2019). Language-dependent knowledge acquisition: Investigating bilingual arithmetic learning. *Bilingualism: Language and Cognition*, *22*(1), 47–57. https://doi.org/10.1017/S1366728917000530

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). *Individual differences in non-verbal number acuity correlate with maths achievement*. *455*, 655–668. https://doi.org/10.1038/nature07246

Haspelmath, M., Dryer, M. S., Gil, D., & Comrie, B. (2005). *The World Atlas of Language Structures*. Oxford Univ. Press.

Hernandez, A. E. (2013). *The bilingual brain*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199828111.001.0001

Hirsh, K. W., Morrison, C. M., Gaset, S., & Carnicer, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition*, *6*(2), 117–128. https://doi.org/10.1017/S136672890300107X

Ifrah, G., & Bellos, D. (2000). *The universal history of numbers: From prehistory to the invention of the computer*. Wiley.

Imbo, I., Vanden Bulcke, C., De Brauwer, J., & Fias, W. (2014). Sixty-four or four-and-sixty? The influence of language and working memory on children's number transcoding. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00313

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277–288. https://doi.org/10.1016/j.actpsy.2007.06.003

Kempert, S., Saalbach, H., & Hardy, I. (2011). Cognitive benefits and costs of bilingualism in elementary school students: The case of mathematical word problems. *Journal of Educational Psychology*, *103*(3), 547–561. https://doi.org/10.1037/a0023619

Klaus, J., & Schriefers, H. (2019). Bilingual Word Production. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the Neuroscience of Multilingualism* (pp. 214–229). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119387725.ch10

Kochari, A. R. (2019). Conducting Web-Based Experiments for Numerical Cognition Research. *Journal of Cognition*, *2*(1), 39. https://doi.org/10.5334/joc.85

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed Numbers: Exploring the Modularity of Numerical Representations With Masked and Unmasked Semantic Priming. *Journal of Experimental Psychology: Human Perception and Performance*, 24. https://doi.org/10.1037/0096-1523.25.6.1882

Kovelman, I., Baker, S. A., & Petitto, L.-A. (2008). Age of first bilingual language exposure as a new window into bilingual reading development. *Bilingualism: Language and Cognition*, *11*(2), 203–223. https://doi.org/10.1017/S1366728908003386

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01795

Krinzinger, H., Gregoire, J., Desoete, A., Kaufmann, L., Nuerk, H.-C., & Willmes, K. (2011). Differential Language Effects on Numerical Skills in Second Grade. *Journal of Cross-Cultural Psychology*, *42*(4), 614–629. https://doi.org/10.1177/0022022111406252

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, *13*(3), 373–381. https://doi.org/10.1017/S136672891000009X

Lachelin, R., Marinova, M., Reynvoet, B., & Schiltz, C. (2023). Weaker semantic priming effects with number words in the second language of math learning. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001526

Lachelin, R., Rinsveld, A. van, Poncin, A., & Schiltz, C. (2022). Number transcoding in bilinguals—A transversal developmental study. *PLOS ONE*, *17*(8), e0273391. https://doi.org/10.1371/journal.pone.0273391

Lê, M.-L., & Noël, M.-P. (2021). Preschoolers' mastery of advanced counting: The best predictor of addition skills 2 years later. *Journal of Experimental Child Psychology*, *212*, 105252. https://doi.org/10.1016/j.jecp.2021.105252

Lê, M.-L. T., & Noël, M.-P. (2020). Transparent number-naming system gives only limited advantage for preschooler's numerical development: Comparisons of Vietnamese and French-speaking children. *PLOS ONE*, *15*(12), e0243472. https://doi.org/10.1371/journal.pone.0243472

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia*, *41*(14), 1942–1958. https://doi.org/10.1016/S0028-3932(03)00123-4

Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means* [Manual]. https://CRAN.R-project.org/package=emmeans

Lin, J.-F. L., Imada, T., & Kuhl, P. K. (2012). Mental Addition in Bilinguals: An fMRI Study of Task-Related and Performance-Related Activation. *Cerebral Cortex*, *22*(8), 1851–1861. https://doi.org/10.1093/cercor/bhr263

Lin, J.-F. L., Imada, T., & Kuhl, P. K. (2019). Neuroplasticity, bilingualism, and mental mathematics: A behavior-MEG study. *Brain and Cognition*, *134*, 122–134. https://doi.org/10.1016/j.bandc.2019.03.006

Lonnemann, J., & Yan, S. (2015). Does number word inversion affect arithmetic processes in adults? *Trends in Neuroscience and Education*, *4*(1), 1–5. https://doi.org/10.1016/j.tine.2015.01.002

Major, C. S., Paul, J. M., & Reeve, R. A. (2017). TEMA and Dot Enumeration Profiles Predict Mental Addition Problem Solving Speed Longitudinally. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02263

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115. https://doi.org/10.1017/S1366728903001068

Marinova, M., Georges, C., Guillaume, M., Reynvoet, B., Schiltz, C., & Van Rinsveld, A. (2021). Automatic integration of numerical formats examined with frequency-tagged EEG. *Scientific Reports*, *11*(1), 21405. https://doi.org/10.1038/s41598-021-00738-0

Marsh, L. G., & Maki, R. H. (1976). Efficiency of arithmetic operations in bilinguals as a function of language. *Memory & Cognition*, *4*(4), 459–464. https://doi.org/10.3758/BF03213203

Martinez-Lincoln, A., Cortinas, C., & Wicha, N. Y. Y. (2015). Arithmetic memory networks established in childhood are changed by experience in adulthood. *Neuroscience Letters*, *0*, 325–330. https://doi.org/10.1016/j.neulet.2014.11.010

Martini, S. (2021). *The influence of language on mathematics in a multilingual educational setting*. University of Luxembourg.

McClain, L., & Huang, J. Y. S. (1982). Speed of simple arithmetic in bilinguals. *Memory & Cognition*, *10*(6), 591–596. https://doi.org/10.3758/BF03202441

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, *44*(1–2), 107–157. https://doi.org/10.1016/0010-0277(92)90052-J

McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition*, *4*(2), 171–196. https://doi.org/10.1016/0278-2626(85)90069-7

McClung, N. A., & Arya, D. J. (2018). Individual Differences in Fourth-Grade Math Achievement in Chinese and English. *Frontiers in Education*, *3*, 29. https://doi.org/10.3389/feduc.2018.00029

Meeuwissen, M., Roelofs, A., & Levelt, W. J. M. (2003). Planning levels in naming and reading complex numerals. *Memory & Cognition*, *31*(8), 1238–1248. https://doi.org/10.3758/BF03195807

Miller, K. F., Smith, C. M., Zhu, J., & Zhang, H. (1995). Preschool Origins of Cross-National Differences in Mathematical Competence: The Role of Number-Naming Systems. *Psychological Science*, *6*(1), 56–60.

Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development*, *2*(3), 279–305. https://doi.org/10.1016/S0885-2014(87)90091-8

Ministère de l'Éducation Nationale. (2022, April 20). Languages in Luxembourg Schools. https://men.public.lu/en/themes-transversaux/langues-ecole-luxembourgeoise.html

Miura, I. T., Kim, C. C., Chang, C.-M., & Okamoto, Y. (1988). Effects of Language Characteristics on Children's Cognitive Representation of Number: Cross-National Comparisons. *Child Development*, *59*(6), 1445. https://doi.org/10.2307/1130659

Moeller, K., Shaki, S., Göbel, S. M., & Nuerk, H.-C. (2015). Language influences number processing – A quadrilingual study. *Cognition*, *136*, 150–155. https://doi.org/10.1016/j.cognition.2014.11.003

Moeller, K., Zuber, J., Olsen, N., Nuerk, H.-C., & Willmes, K. (2015). Intransparent German number words complicate transcoding – a translingual comparison with Japanese. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00740

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), Article 5109. https://doi.org/10.1038/2151519a0

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious Masked Priming Depends on Temporal Attention. *Psychological Science*, *13*(5), 416–424. https://doi.org/10.1111/1467-9280.00474

Naccache, L., & Dehaene, S. (2001). The Priming Method: Imaging Unconscious Repetition Priming Reveals an Abstract Representation of Number in the Parietal Lobes. *Cerebral Cortex*, *11*(10), 966–974. https://doi.org/10.1093/cercor/11.10.966

Negen, J., & Sarnecka, B. W. (2009). *Young children's number-word knowledge predicts their performance on a nonlinguistic number task*. https://escholarship.org/uc/item/1q03q75z

Negen, J., & Sarnecka, B. W. (2012). Number-Concept Acquisition and General Vocabulary Development. *Child Development*, *83*(6), 2019–2027. https://doi.org/10.1111/j.1467-8624.2012.01815.x

Notebaert, K., Pesenti, M., & Reynvoet, B. (2010). The neural origin of the priming distance effect: Distance-dependent recovery of parietal activation using symbolic magnitudes. *Human Brain Mapping*, *31*(5), 669–677. https://doi.org/10.1002/hbm.20896

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition*, *111*(2), 275–279. https://doi.org/10.1016/j.cognition.2009.02.002

Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition*, *82*(1), B25–B33. https://doi.org/10.1016/S0010-0277(01)00142-1

Nuerk, H.-C., Weger, U., & Willmes, K. (2004). On the Perceptual Generality of the Unit-Decade Compatibility Effect. *Experimental Psychology*, *51*(1), 72–79. https://doi.org/10.1027/1618-3169.51.1.72

Nuerk, H.-C., Weger, U., & Willmes, K. (2005). Language effects in magnitude comparison: Small, but not irrelevant. *Brain and Language*, *92*(3), 262–277. https://doi.org/10.1016/j.bandl.2004.06.107

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, *15*(2), 202–206. https://doi.org/10.1016/j.conb.2005.03.007

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503. https://doi.org/10.1126/science.1102085

Pitt, B., Gibson, E., & Piantadosi, S. T. (2022). *Exact number concepts are limited to the verbal count range*. *33*(3), 371–381. https://doi.org/10.1177/09567976211034502

Pixner, S., Zuber, J., Heřmanová, V., Kaufmann, L., Nuerk, H.-C., & Moeller, K. (2011). One language, two number-word systems and many problems: Numerical cognition in the Czech language. *Research in Developmental Disabilities*, *32*(6), 2683–2689. https://doi.org/10.1016/j.ridd.2011.06.004

Poncin, A., Rinsveld, A. V., & Schiltz, C. (2020). *Units first or tens first: How bilingualism affects two-digit number transcoding?* PsyArXiv. https://doi.org/10.31234/osf.io/sg7ea

Poncin, A., Van Rinsveld, A., & Schiltz, C. (2019). Units-first or tens-first: Does language matter when processing visually presented two-digit numbers? *Quarterly Journal of Experimental Psychology*, *73*(5), 726–738. https://doi.org/10.1177/1747021819892165

Power, R. J. D., & Dal Martello, M. F. (1990). The dictation of Italian numerals. *Language and Cognitive Processes*, *5*(3), 237–254. https://doi.org/10.1080/01690969008402106

Prior, A., Katz, M., Mahajna, I., & Rubinsten, O. (2015). Number word structure in first and second language influences arithmetic skills. *Frontiers in Psychology*, *6*(MAR), 266. https://doi.org/10.3389/fpsyg.2015.00266

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859–862. https://doi.org/10.3758/BF03192979

R Core Team. (2013). *R: A language and environment for statistical computing*. http://www.R-project.org/

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, *72*(2), 191–201. https://doi.org/10.1016/S0010-0277(99)00048-7

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic priming in number naming. *The Quarterly Journal of Experimental Psychology Section A*, *55*(4), 1127–1139. https://doi.org/10.1080/02724980244000116

Reynvoet, B., De Smedt, B., & Van den Bussche, E. (2009). Children's representation of symbolic magnitude: The development of the priming distance effect. *Journal of Experimental Child Psychology*, *103*(4), 480–489. https://doi.org/10.1016/j.jecp.2009.01.007

Rodic, M., Zhou, X., Tikhomirova, T., Wei, W., Malykh, S., Ismatulina, V., Sabirova, E., Davidova, Y., Tosto, M. G., Lemelin, J.-P., & Kovas, Y. (2015). Cross-cultural investigation into cognitive underpinnings of individual differences in early arithmetic. *Developmental Science*, *18*(1), 165–174. https://doi.org/10.1111/desc.12204

RStudio Team. (2020). *RStudio: Integrated development environment for r* [Manual]. http://www.rstudio.com/

Saalbach, H., Eckstein, D., Andri, N., Hobi, R., & Grabner, R. H. (2013). When language of instruction and language of application differ: Cognitive costs of bilingual mathematics learning. *Learning and Instruction*, *26*, 36–44. https://doi.org/10.1016/j.learninstruc.2013.01.002

Salillas, E., Barraza, P., & Carreiras, M. (2015). Oscillatory Brain Activity Reveals Linguistic Prints in the Quantity Code. *PLOS ONE*, *10*(4), e0121434. https://doi.org/10.1371/journal.pone.0121434

Salillas, E., & Carreiras, M. (2014). Core number representations are shaped by language. *Cortex*, *52*, 1–11. https://doi.org/10.1016/j.cortex.2013.12.009

Salillas, E., & Martínez, A. (2018). Linguistic Traces in Core Numerical Knowledge: An Approach From Bilingualism. In *Language and Culture in Mathematical Cognition* (pp. 173–196). Elsevier. https://linkinghub.elsevier.com/retrieve/pii/B9780128125748000080

Salillas, E., & Wicha, N. Y. Y. (2012). Early Learning Shapes the Memory Networks for Arithmetic: Evidence From Brain Potentials in Bilinguals. *Psychological Science*, *23*(7), 745–755. https://doi.org/10.1177/0956797612446347

Sasanguie, D., Defever, E., Van den Bussche, E., & Reynvoet, B. (2011). The reliability of and the relation between non-symbolic numerical distance effects in comparison, same-different judgments and priming. *Acta Psychologica*, *136*(1), 73–80. https://doi.org/10.1016/j.actpsy.2010.10.004

Sasanguie, D., & Reynvoet, B. (2014). Adults' Arithmetic Builds on Fast and Automatic Processing of Arabic Digits: Evidence from an Audiovisual Matching Paradigm. *PLOS ONE*, *9*(2), e87739. https://doi.org/10.1371/journal.pone.0087739

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., & Smedt, B. D. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3), e12372. https://doi.org/10.1111/desc.12372

Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology*, *12*(3), 281–300. https://doi.org/10.1111/j.2044-835X.1994.tb00635.x

Singmann, H. (2021). *Mixed Model Reanalysis of RT data* [Computer software]. https://cran.r-project.org/web/packages/afex/vignettes/afex_mixed_example.html

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *afex: Analysis of factorial experiments* [Computer software]. https://CRAN.R-project.org/package=afex

Spaepen, E., Coppola, M., Flaherty, M., Spelke, E., & Goldin-Meadow, S. (2013). Generating a lexicon without a language model: Do words for number count? *Journal of Memory and Language*, *69*(4), 496–505. https://doi.org/10.1016/j.jml.2013.05.004

Spelke, E. S., & Tsivkin, S. (2001a). Initial knowledge and conceptual change: Space and number. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (pp. 70–98). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.005

Spelke, E. S., & Tsivkin, S. (2001b). Language and number: A bilingual training study. *Cognition*, *78*(1), 45–88. https://doi.org/10.1016/S0010-0277(00)00108-6

Steiner, A. F., Banfi, C., Finke, S., Kemény, F., Clayton, F. J., Göbel, S. M., & Landerl, K. (2021). Twenty-four or four-and-twenty: Language modulates cross-modal matching for multidigit numbers in children and adults. *Journal of Experimental Child Psychology*, *202*, 104970. https://doi.org/10.1016/j.jecp.2020.104970

Steiner, A. F., Finke, S., Clayton, F. J., Banfi, C., Kemény, F., Göbel, S. M., & Landerl, K. (2021). Language effects in early development of number writing and reading. *Journal of Numerical Cognition*, *7*(3), 368–387. https://doi.org/10.5964/jnc.6929

Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual Word Recognition in a Sentence Context. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00174

van der Ven, S. H. G., Klaiber, J. D., & van der Maas, H. L. J. (2017). Four and twenty blackbirds: How transcoding ability mediates the relationship between visuospatial working memory and math in a language with inversion. *Educational Psychology*, *37*(4), 487–505. https://doi.org/10.1080/01443410.2016.1150421

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

van Marle, K., Chu, F. W., Li, Y., & Geary, D. C. (2014). Acuity of the approximate number system and preschoolers' quantitative development. *Developmental Science*, *17*(4), 492–505. https://doi.org/10.1111/desc.12143

Van Rinsveld, A., Brunner, M., Landerl, K., Schiltz, C., & Ugen, S. (2015). The relation between language and arithmetic in bilinguals: Insights from different stages of language acquisition. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00265

Van Rinsveld, A., Dricot, L., Guillaume, M., Rossion, B., & Schiltz, C. (2017). Mental arithmetic in the bilingual brain: Language matters. *Neuropsychologia*, *101*, 17–29. https://doi.org/10.1016/j.neuropsychologia.2017.05.009

Van Rinsveld, A., & Schiltz, C. (2016). Sixty-twelve = Seventy-two? A cross-linguistic comparison of children's number transcoding. *British Journal of Developmental Psychology*, *34*(3), 461–468. https://doi.org/10.1111/bjdp.12151

Van Rinsveld, A., Schiltz, C., Brunner, M., Landerl, K., & Ugen, S. (2016). Solving arithmetic problems in first and second language: Does the language context matter? *Learning and Instruction*, *42*, 72–82. https://doi.org/10.1016/j.learninstruc.2016.01.003

Van Rinsveld, A., Schiltz, C., Landerl, K., Brunner, M., & Ugen, S. (2016). Speaking two languages with different number naming systems: What implications for magnitude judgments in bilinguals at different stages of language acquisition? *Cognitive Processing*, *17*(3), 225–241. https://doi.org/10.1007/s10339-016-0762-9

Vander Beken, H., & Brysbaert, M. (2018). Studying texts in a second language: The importance of test type. *Bilingualism: Language and Cognition*, *21*(5), 1062–1074. https://doi.org/10.1017/S1366728917000189

Venkatraman, V., Siong, S. C., Chee, M. W. L., & Ansari, D. (2006). Effect of Language Switching on Arithmetic: A Bilingual fMRI Study. *Journal of Cognitive Neuroscience*, *18*(1), 64–74. https://doi.org/10.1162/089892906775250030

Volmer, E., Grabner, R. H., & Saalbach, H. (2018). Language switching costs in bilingual mathematics learning: Transfer effects and individual differences. *Zeitschrift Für Erziehungswissenschaft*, *21*(1), 71–96. https://doi.org/10.1007/s11618-017-0795-6

Wang, Y., Lin, L., Kuhl, P., & Hirsch, J. (2007). Mathematical and Linguistic Processing Differs Between Native and Second Languages: An fMRI Study. *Brain Imaging and Behavior*, *1*(3–4), 68–82. https://doi.org/10.1007/s11682-007-9007-y

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, *8*(3), 231–256. https://doi.org/10.1162/jocn.1996.8.3.231

Wicha, N. Y., Dickson, D. S., & Martinez-Lincoln, A. (2018). Arithmetic in the Bilingual Brain. In *Language and Culture in Mathematical Cognition* (pp. 145–172). Elsevier. https://doi.org/10.1016/B978-0-12-812574-8.00007-9

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251. https://doi.org/10.1016/0010-0285(92)90008-P

Xenidou-Dervou, I., Gilmore, C., van der Schoot, M., & van Lieshout, E. C. D. M. (2015). The developmental onset of symbolic approximation: Beyond nonsymbolic representations, the language of numbers matters. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00487

Xenidou-Dervou, I., van Atteveldt, N., Surducan, I. M., Reynvoet, B., Rossi, S., & Gilmore, C. (2023). Multiple number-naming associations: How the inversion property affects adults' two-digit number processing. *Quarterly Journal of Experimental Psychology*, 17470218231181367. https://doi.org/10.1177/17470218231181367

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Ziegler, J. C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical Words are Read Differently in Different Languages. *Psychological Science*, *12*(5), 379–384. https://doi.org/10.1111/1467-9280.00370

Zuber, J., Pixner, S., Moeller, K., & Nuerk, H.-C. (2009). On the language specificity of basic number processing: Transcoding in a language with inversion and its relation to working memory capacity. *Journal of Experimental Child Psychology*, *102*(1), 60–77. *https://doi.org/10.1016/j.jecp.2008.04.003*

1 STUDY 2

Units first or Tens first: How bilingualism affects two-digit auditory-visual number matching

Rémy Lachelin[1], Alexandre Poncin[1], Amandine Van Rinsveld[2], & Christine Schiltz[1]
[1] University of Luxembourg, [2] Université libre de Bruxelles

Status: submitted

Pre-print: https://osf.io/preprints/psyarxiv/cdyx4

Author note

The authors made the following contributions. Rémy Lachelin: Writing - Original Draft Preparation, Data Analyses, Writing - Review & Editing; Alexandre Poncin: Conceptualization, Pre-registration, Writing - Original Draft Preparation; Amandine Van Rinsveld: Conceptualization, Pre-registration; Christine Schiltz: Conceptualization, Pre-registration, Writing - Review & Editing.

1.1 Abstract

In German the number word for *4*2 is inverted such that "*forty*-two" is "two-and-*forty*", while this is not the case in French. German and French monolingual and bilingual speakers' number processing might be impacted by this inversion between Arabic digits and number word formats. We compared German-French bilinguals who sequentially learned mathematics first in German (LM1, until 7^{h} grade) and then in French (LM2 until the end of secondary school) with language matched monolinguals. In an auditory-visual number matching task participants heard a number word which had to be matched with a visually presented Arabic number among three distractors. In two additional sequential conditions, we first presented tens (4_) or units (_2) before presenting the entire number (42). While we did not observe performance differences between both monolingual groups, bilinguals displayed a similar behavioural pattern in all conditions in their LM1 (German). But they were overall slightly slower than German monolinguals, hence displaying a bilingual lexical cost. Bilinguals were also impacted by a global LM2 processing cost, being overall slower in the LM2 (French) than the LM1. Furthermore, in the LM2, the conditions interacted with language. On the one hand relatively enhanced interference from the Unit-first condition performed in LM2 is suggesting an influence from LM1's inverted morpho-syntactic structure. On the other hand an increased facilitation in the Ten-first condition indicates a relative enhanced facilitation when the LM2 is transparent with regards to Arabic number's positional order. These interference effects suggest that bilingual's processing strategies can flexibly transfer across LM1 and LM2 and adapt to the morpho-syntactic features of their spoken languages.

Keywords: Bilinguals, language of learning mathematics, number word inversion, number auditory visual matching task

2 Introduction

Symbolic numbers such as Arabic numbers (42) or number words (forty-two), are required for precise mental representation of large numbers (Lemer et al., 2003; Pica et al., 2004). Symbolic number systems have evolved with human cultures, hence some of them have been adopted across different languages and cultures. The Indo-Arabic system is a base-10 place-value system in use across many languages. Conversely, number words, more attached to oral transmission, are language dependent (Comrie, 2013). For example, in French 42 is named starting by the number of tens and then units “quarante-deux”, but in German naming starts by the unit digit and then proceeds to the ten digit “zwei-und-vierzig” (Ifrah & Bellos, 2000). Language morpho-syntactic differences affect number learning and mathematics (*i.e.* Dowker & Nuerk, 2016). Several cognitive models describe the processing of morpho-syntactic characteristics (Barrouillet et al., 2004; Dehaene, 1992; Dotan & Friedmann, 2018). However, despite more than half of the word population being bilingual (Grosjean, 2010) few models and studies integrate multiple languages into their description of number processing (some examples: Bernardo, 2005; Campbell & Epp, 2004). This pre-registered study aimed to address this shortcoming by investigating how bilingual adults process two-digit numbers. The bilinguals learned mathematics first in German and then in French; we compared them to German and French monolinguals in an auditory-visual matching task.

2.1 Language-dependent number word structures

Transparency is the degree of morpho-syntactic correspondence between visual and verbal numbers and is therefore language-dependent. For example, languages such as German or French have morpho-syntactic characteristics that make them opaquer relative to Indo-Arabic numbers. Importantly, lack of transparency can interfere and delay arithmetic and number acquisition.

For German and Dutch numbers for example, the ten-unit place-value order position is inverted compared to visual Arabic digits (i.e. *4*2 is “Zwei und *Vierzig*”, literally “two and *forty*”), this inversion hence affects the transparency of order (Bahnmueller et al., 2018). German speaking children are slower compared to their Italian speaking peers for mathematical problems involving a change of tens requiring a carry operation, i.e. in 28 + 16 (Göbel, Moeller, et al., 2014). This carry effect, doing more errors in problems needing a carry, is also found in German speaking adults in comparison to Chinese speakers (Lonnemann & Yan, 2015). In a comparison of English and Dutch speaking 5 year-old children, Xenidou-Dervou et al., (2015) found that Dutch speakers were delayed with regards to arithmetic performances. Furthermore, arithmetic performances correlated with number naming, suggesting a relation between inversion and delayed arithmetic skill acquisition.

Further experimental studies also showed an impact of inversion on more basic numerical tasks, such as judging which of two numbers is the largest. In number comparison tasks, faster responses and lower error rates are found in adults for compatible trials in which both tens and units lead to the same decision (e.g.: 42 < 87 , 4 < 8 and 2 < 7) than incompatible trials (e.g. : 28 < 42, 2 < 4 but 8 > 2) (Nuerk et al., 2001, 2004). The compatibility effect is larger with Arabic numbers in units (i.e. 42 < 47) for German speakers than speakers of non-inverted languages such as English (Nuerk et al., 2005). The compatibility effect is also found in German-French bilinguals, independently from the increasing proficiency in French (Van Rinsveld, Schiltz, Landerl, et al., 2016).

The difficulty with inversion further generalizes from arithmetic and number comparison to less cognitively demanding tasks involving number words, especially for children. These are also called transcoding tasks since they involve the conversion between numerical codes (e.g., a number dictation tasks involves the conversion from verbal to visual). In a number writing task done by German speaking first graders, about half of the errors could be explained by inversion (Zuber et al., 2009). First and second grade Dutch speakers, made more inversion errors than French speakers in number dictation tasks (Imbo et al., 2014). 8 years-old German speaking children also made more inversion-related errors in writing and reading numbers than English-speaking matched peers (Steiner, Finke, et al., 2021). German speaking children do more syntactic inversion errors than Japanese speaking peers when writing down numbers (Moeller, Zuber, et al., 2015). Some authors have, however, noted that those differences might at least partly be explained by curricular differences between countries (Krinzinger et al., 2011). Still, more inversion errors are also found using within subject designs, hence within the same school curriculum. In the Czech language, ten-unit and unit-ten order for number words co-exist and Czech speaking children doing a number dictation task do more errors in the inverted unit-ten than the non-inverted ten-unit number word inversion (Pixner et al., 2011). In a large sample of Dutch speaking children, van der Ven et al., (2017) found that for 6th grade and younger children more than half of the children did at least one inversion error in a series of web-based number transcoding games. In sum, transparency of order, i.e., the inversion of tens and units in some languages, influences arithmetic, magnitude comparison, and transcoding tasks.

Number matching tasks have also been used to experimentally test the effect of ten-unit inversion. In these tasks a number is presented orally followed by visually presented Arabic digits, hence requiring the participant to transcode an auditory number onto a visual one. Participant's task is either to respond if the two numbers match (are the same) or to find the matching number among distractors. Number matching tasks predict arithmetic performances, suggesting overlapping cognitive processes for both tasks (Sasanguie & Reynvoet, 2014). Steiner, Banfi, et al., (2021) compared German to English speaking 2nd, 3rd grade children (longitudinal) and adults with a two-digit number matching task. Half of the trials were matching Arabic numbers, while in the other half one of four non-matching categories of distractors were presented. For example, for the target "*twenty*-four" in English but "vier-und-*zwanzig*" in German (24) the non-matching stimuli were: inversions distractor (42), unit distractors (28), ten distractor (48) and unrelated distractor (36). German speaking children and adults were slower to reject inversion distractors compared to unit distractors, while no such pattern could be found in English speaking adults. Therefore, suggesting an interference in adult German monolinguals when the distractors are congruent with the ten-unit order but not matching the quantity (i.e. "vier-und-*zwanzig*" ≠ *4*2). In a study that is partially replicated and extended here, Poncin et al., (2019) investigated how German and French monolingual children and adults are affected in matching auditory to visual two-digit number (i.e. 42) when preceded by either the unit (_2) or the ten (4_) part of the number (i.e. sequential conditions). For adults there were no performance differences between the sequential conditions. However, for children, French monolinguals were faster in the Ten-first than the Unit-first condition, while no difference between condition was observed in German. In sum, number matching tasks confirm the effect of inversion on Arabic numbers in children and adults speaking a language with inverted number words.

2.2 Bilingual number processing

In addition to language-dependent factors such as the inversion in German, individual language profiles might influence numerical cognition, especially for bilinguals. Given that about half of the global population is bilingual (Grosjean, 2010), it is important to assess how learning and consolidating languages, especially through formal education, affects numerical cognition in bilinguals[13]. Yet, bilingual's language profiles can be very heterogeneous, for example bilingual can be balanced or unbalanced and have high or low proficiencies. Highly proficient balanced bilinguals have comparable and high proficiency in both languages (de Groot, 2011). Age, or order of language acquisition, can also affect balance and proficiency, with higher proficiencies for the first learned language (L1) than the second language (L2). In the brain, the effect of language dominance and proficiency is linked to top-down frontal inhibitory mechanisms: since L1 and L2 are always activated the L2 is easier to inhibit compared to the L1, leading to a facilitation for L1 (Abutalebi, 2008; Green, 1998). Since language influences how arithmetic and numbers are processed, the specific language profile of bilinguals also impacts number processing in both of their languages. The inhibition of co-activated languages of bilinguals, means there might be an additional cognitive process compared to monolinguals, resulting in a general bilingual lexical retrieval cost (Ivanova & Costa, 2008), (Bylund et al., 2023).

Language of formal instruction of mathematics at school, hereafter the language of math acquisition (LM) influences arithmetic and transcoding performances. Philippino home language (HL) students who learned mathematics in English at school (LM+ = English), performed better doing arithmetic in English than in Philippino (Bernardo, 2001). Thus suggesting the importance of LM over HL for arithmetic performances. In an experiment involving participants trained for arithmetic facts in Russian or English, the results showed better performances for exact calculations in the trained than untrained language, this independently from the language dominance (Spelke & Tsivkin, 2001b). Other authors have described this as a language switching cost (LSC): a cost when being tested in a different language than the language of training (Grabner et al., 2012a; Hahn et al., 2017, 2019; Kempert et al., 2011; Saalbach et al., 2013; Volmer et al., 2018). Salillas and Wicha, (2012) showed that participants' brain potentials resulting from the presentation of simple multiplications (i.e., multiplication tables) in LM+ were similar to brain potentials obtained from simple digit presentation, while the presentation of simple multiplication in LM- led to qualitatively different brain potentials. This finding was also true when LM+ was not the dominant language. These results are in line with fMRI studies showing larger activation when doing arithmetic in the L2 compared to L1, hence suggesting a higher cognitive demand in L(M)2 (Lin et al., 2012;

[13] For conciseness and clarity, we only discuss and analyze the specific simpler case of bilingualism as the encompassing case of multilingualism.

Van Rinsveld et al., 2017). In Luxembourg, bilinguals sequentially learn mathematics first in German (LM1) and then in French (LM2). Several studies have found an LM2 cost (i.e. slower responses) when comparing both languages for solving simple and complex arithmetic (Van Rinsveld et al., 2015), two-digit number matching (Lachelin et al., 2022) and single-digit number reading (Lachelin et al., 2023). The LM2 cost is more broadly confirmed for the L2 by meta-analysis (Garcia et al., 2021). In sum, these studies show the effect of language on bilingual's numerical skills: in general bilinguals have better performances in the language they have learned and consolidated mathematics first.

Language profiles and morpho-syntactic language properties can interact in bilinguals. Arabic (L1) - Hebrew (L2) bilinguals were presented auditorily and visually with simple arithmetic problems starting by the units (5 + 20), which matches the (inverted) order in Arabic or the tens (20 + 5), matching Hebrew's number word morpho-syntax. When presented auditorily, the bilinguals were more accurate with the unit first display in Arabic and ten first in Hebrew. When presented visually, bilinguals did not differ among both languages (Prior et al., 2015b). Therefore suggesting these bilinguals were able to flexibly take advantage of the L2 non inverted morpho-syntax in auditory presentation modes. In an adapted version of the auditory-visual number matching task Xenidou-Dervou et al., (2023) manipulated the auditory presented numbers, rather than the visual, by creating artificial numbers in Dutch. Four categories of artificial numbers were created to present either a matching/unmatching quantity and/or being congruent with Arabic number's ten-unit order. Taking 42 as example, both "forty and two" and "two and forty" matched quantity, yet only "two and forty" is incongruent with ten-unit order of Arabic digits. "Two and forty" on the other side, is an existing Dutch word, matching the quantity 42. Two additional artificial words were created with non-matching quantity: "four and twenty" and "twenty and four", where "four and twenty" is congruent with the Arabic order (42). "Four and twenty" is also an existing number word in Dutch. Dutch (L1)-English (L2) bilinguals were less accurate rejecting the number word for the Dutch artificial number words corresponding to "forty and two" than the traditional Dutch number word "two and forty". Moreover, these rejection errors were linked with English proficiency, suggesting that the L2 morpho-syntactic structure influences number processing in L1. In sum, the morpho-syntax of both languages of a bilingual can interact and these between-language transfers might depend on the specific language profile of bilinguals.

2.3 Bilingual Triple Code Model

The triple code model (TCM, Dehaene, 1992) stipulates the existence of three different codes to represent numbers: an analogic code which is related to quantity (*i.e.* the number of visible craters on the moon) a symbolic visual code (*i.e.* Arabic numbers: *4*2) and a verbal code (*i.e.* number words: "*forty*-two"). The three codes are associated by three different routes, allowing to pass from a code to another, a process which is also called transcoding. Bilinguals, however, can transcode from - and to - two languages. We hence stipulated separate language-dependent routes for each language into a bilingual triple code model (BTCM, Lachelin et al., 2023). For example, in the case of German-French bilinguals, *4*2 has two verbal correspondents: in German "Zwei-und-*Vierzig*" (inverted) and in French "*Quarante*-deux" (non-inverted). In the BTCM, the consolidation of each language-dependent route affects the strength of association for each language. The consolidation of each language is the result of subject-dependent language profiles such as balanced or unbalanced proficiency in both languages, age of second language acquisition or number word acquisition, consolidation mostly depending on the language of mathematical education and morpho-syntactic properties of the two languages. In the specific case of German-French bilinguals for instance, one might thus expect interfering effects on the verbal-visual code associations when the morpho-syntax differs from visual numbers (*i.e.,* in German, which is less transparent due to number word inversion). In contrast, their correspondence should lead to facilitation effects (*i.e.,* in French, which is non-inverted).

2.4 Present study

In the present study, we assessed transcoding performances and the underlying cognitive mechanisms of German-French bilingual adults and monolingual German- and French-speaking peers. As in (Poncin et al., 2019) we used a transcoding paradigm in which participants had to listen to two-digit numbers and match them with a visually presented target stimulus presented among three distractors. Manipulating the order of appearance of tens and units led to three priming conditions: *Ten-first condition*, in which the tens of the target and the three distractors appeared before the units (4_ → 42). This condition mimics the French number-word structure, with tens being pronounced before units (e.g. "quarante-deux" corresponding to "forty-two"). *Unit-first condition*, in which the units of the target and the three distractors appeared before the tens (_2 → 42). This condition mimics the German number-word system, with units being pronounced before tens (e.g. "Vier-und-Zwanzig" corresponding to "four-and-twenty"). *Simultaneous condition*, in which tens and units appeared at the same time (42). This condition is the more natural and ecological one, as in everyday life we are typically confronted with two-digit numbers in this format.

2.5 Hypothesis

The following hypotheses were pre-registered before bilingual and monolingual German data was collected (*https://osf.io/b4p2z/*). Because each condition might be influenced by several opposing effects, we set the hypotheses with different effects and directions, see H(A) to H(F). We then calculated a model to make numerical predictions of reaction time for the different conditions in the different language profiles, see F(A). The initial reference for this model (i.e. the intercept, or when all other terms are = 0) is the monolingual French in the simultaneous condition.

$$F(A){:}\ RT = Intercept + Prime + Inversion + Ten\ first \mid Language + Unit\ first|Language + Bilingual\ lexical\ cost + Bilinguals\ LM2\ cost$$

Hereafters are the different hypotheses for each part of F(A):

H (A) *"Prime*": medium facilitation (-60 ms) for all sequential conditions (Ten-first and Unit-first) compared to the simultaneous condition. We predicted a facilitation for sequential compared to the simultaneous conditions due to the availability of ten or unit information before the response.

H (B) *Inversion*: small hindering (30 ms) for all conditions in German. This hypothesis regards the *language-dependent* verbal stimuli presentation: the inversion of number words in German compared to the non-inverted number word structure in French should have a small effect somewhat slowing down responses in German.

H (C) *Ten-first* (i.e. 4_) condition: large facilitation in French (-100 ms), but only small facilitation in German (-30 ms). Since this condition mimics the French number word system starting with tens, we expected a larger facilitation in French than German. Nevertheless, some facilitation is also expected in German, due to the congruence of this presentation format with the Ten-first order position of Arabic numbers (see Xenidou-Dervou et al., 2023).

H (D) *Unit-first* (i.e. _2): small interference in French (30 ms), but medium facilitation in German (-60 ms). The Unit-first condition was designed to mimic the German inverted number word system, while it is incongruent with the French number word order. We therefore expect two opposite effects: a facilitation in German, but a hindering in French due to interference from an unexpected order of presentation.

Hypotheses H(E) & H(F) concern the *subject-dependent* language profile; hence those apply only to the bilingual group. H(F) is specific for the task performed by bilinguals in French.:

H (E) *Bilingual lexical cost* compared to monolinguals in all conditions: medium hindering for bilinguals (in German and French) compared to monolinguals (60 ms). A generally slower lexical retrieval in bilinguals compared to monolinguals is expected (i.e. bilingual lexical cost) due to possible lexical competition between languages when bilinguals are seeing an Arabic number.

H (F) *Bilingual LM2 cost* arising when the bilingual group is doing the task in French (LM2) compared to German (60ms). Since bilinguals learned mathematics first in German (LM1) and then in French (LM2), we expect an LM2 cost, such as found in previous studies with this specific bilingual profile (i.e. Lachelin et al., 2023; Van Rinsveld et al., 2015).

Using the above hypothesis (H(A) to H(F)) applied to F (A) with a predefined intercept for the monolingual French group in the simultaneous condition, we calculated a numerical predictions of reaction times in all cases of different groups and conditions in Table in SM 1.

3 Methods

3.1 Population

German and French bilingual adults were recruited by email and from the internal website among the student population of the university of Luxembourg. The recruitment flyer specified participants must speak both German and French and have spent at least 10 years of schooling in Luxembourg. The two monolingual groups were recruited with prolific (www.prolific.com), the criteria were: being between 18 and 25 years old and being either German or French native speakers. We first recruited 40 participants as in the pre-registration, however this initial sample was not balanced for gender, hence we recruited additional 20 participants for each group. The required sample size was determined by a power analysis based on previously collected data (Poncin et al., 2020), and is available here: *https://osf.io/b4p2z/*. The power analyses were made using simulations of linear mixed models.

The following exclusion criterium were applied per participants. From the 265 participants who completed all three conditions, we excluded: 15 participants who completed less than 80 of the 90 trials in one of the conditions, 6 participants with more than 10 % trial errors on the main task (extreme value), 1 participant with a very low TTR score (i.e. less than 30 operations resolved) and 2 participants without TTR scores, 4 participants older than 30 which was our inclusion criterium: thus leading to a final sample of 237. From this intermediary sample, we excluded 4 monolingual French who did not report French as L1, 2 monolingual German who did not report German as L1 and 2 monolingual German with French as L2 (no German L2 were found in the monolingual French). In the bilingual group we excluded 5 participant who did the experiment only in one language, 2 who did not report speaking German nor French, 14 who reported French as their most proficient language, 17 who did not report speaking Luxembourgish as the L1 and finally 30 who reported Portuguese as L1 or L2. The exclusion of native French speakers is justified since we want to measure the effect of LM2. The exclusion of Portuguese speakers from the analyses was justified because they generally score lower in reading and mathematics scores compared to Luxembourgish speakers, likely due to burden of mastering German and French (and Luxembourgish) in addition to Portuguese for school (Greisen et al., 2021; Martini, 2021). Hence the final sample is of: N =161, see Table 3 for descriptive information/demographics.

Table 3: Sample demographics.

Group		Mono. German	Mono. French	Bilinguals
N(*Women*)		55(*28*)	56(*24*)	50(*32*)
Mean Age		22.2	22.9	21.9
N Languages		2.82	2.73	4.64
TTR(*SD*)		114.22(*20*)	106.11(*21*)	111.32(*21*)
AoA	German	1.16 (6 max)	.	4.56 (7 *max*)
	French	.	1.32 (16 max)	6.82 (10 *max*)
	Lux.	.	.	1.54 (6 max)
AoA Math	German	5.80 (8 max)	.	5.28 (12 max)
	French	.	5.28 (14 max)	7.96 (15 max)
Frequency	German	D54-W1-M0-Y0-N0	D0-W1-M1-Y6-N2	D17-W25-M6-Y1-N0
	French	D0-W2-M2-Y4-N0	D54-W2-M0-Y0-N0	D17-W27-M6-Y0-N0

Notes: "N Languages" is the average number of languages the participants reported speaking. All three groups had comparable TTR scores as revealed by an ANOVA (F(2,158) = 2.03, p = 134). AoA = Age of Acquisition. For frequency: D = Day, W = Weeks, M = Months, Y = Year, N = Never

Monolinguals reported acquiring their L1 and learning mathematics (LM) at about the same age. Bilinguals spoke in average 2 more languages than monolinguals: Luxembourgish, and English for the large majority. The reported L1 Luxemburgish, is linguistically as close to German as other German dialects are (Martini, 2021). All bilinguals attended the same school system where primary school starts at the age of 6 years in German as a general instruction language and specifically for mathematics, which roughly corresponds to the 5.25 years reported, i.e. LM1. French is learned from second grade in primary school as a second language with about 7 years. At age 12, students start secondary school (composed of 7 grades) and French then also becomes the instruction language for mathematics, i.e. LM2 (*Ministère de l'Éducation Nationale*, 2022). When reaching the highest grades of secondary school participants were thus bilinguals with high proficiency levels in both German and French, corresponding to a level of proficiency equivalent to C1 in terms of European Framework of Reference. Hence, the present sample corresponds to German-French bilinguals with similar trajectories not only of language acquisition but also of language of instruction for mathematics.

3.2 *Ethical concerns*

Consent was requested before starting the experiment. The local Ethics Review Panel approved the study (ERP 22-067 NMBilGF). Bilingual participants were rewarded by a voucher usable online, while "prolific platform" participants were rewarded through the platform's services.

3.3 *Materials and Procedures*

Both the number matching task and Tempo Test Rekennen (TTR, De Vos, 1992) were encoded in Labvanced (Finger et al., 2017). The number matching task consisted in listening to a spoken number word and finding the matching Arabic number amongst four numbers presented on a computer screen (see Figure 12). Critically, we constructed three conditions in which we systematically manipulated the timing at which unit- and ten-digits appeared on the screen. In the ecological "simultaneous" condition tens and units appeared simultaneously. In the two sequential presentation conditions ten-digits appeared 500ms before the unit-digits in the Ten-first condition and unit-digits appeared 500ms before the ten-digits in the Unit-first condition. All three conditions contained the 42 target stimuli and their corresponding three distractors in random order. The order of the three conditions was counterbalanced across participants. Each condition was preceded by 12 warmup items with feedback. Response times are measured from to the period between the onset of the last digit (i.e., ten-digit or unit-digit for the Unit-first and the Ten-first conditions, respectively) and the onset of participants' response. The response-times of sequential and "simultaneous" conditions cannot be compared directly due to differences in information content at the start of the response recording period. While participants need to process the two digits in the "simultaneous" condition, one of the digits (i.e., ten-digit in the Ten-first; unit-digit in the Unit-first) has already been processed in the sequential conditions.

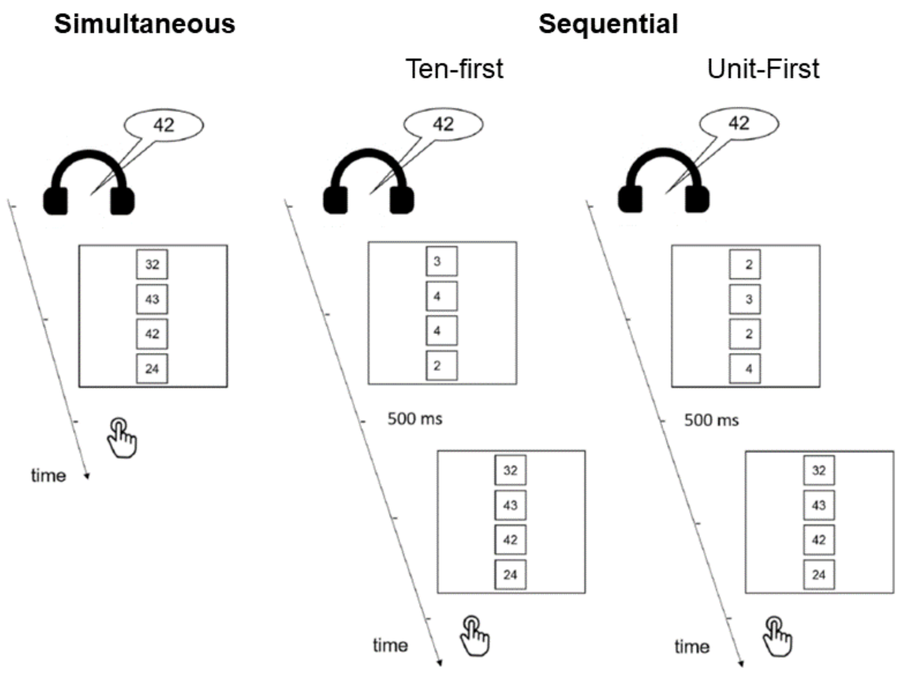


Figure 1: Illustration of the auditory-visual number matching task. Numbers were auditorily presented either in German or in French to bilinguals and monolinguals.

The 90 target-stimuli consisted of all numbers between 23 and 70, except ties (e.g., 22) and tens (e.g., 20). We avoided using numbers from 70 to 90 since those are constructed differently in French (i.e. "seventy" is literally said "sixty and ten", soixante-dix) and would lead to slower reaction times (Lachelin et al., 2022). The three different types of distractor-stimuli were built from the targets: (1) "unit distractors" in which one unit was randomly added or subtracted to the heard number (e.g. for 42 distractor was 43 or 41), (2) "ten distractors" in which ten and units were randomly added or subtracted to the heard number (e.g. for 42 the distractor was 32 or 52), (3) "inversion distractors" where units and tens were inverted (e.g. for 42 the distractor was 24). To avoid bias in our distractor stimuli (i.e., impossibility to create "unit distractors" and "ten distractors" with 1 or 9 while respecting the above-mentioned principles), we also removed numbers containing 1 or 9 (e.g., 31; 92; 49) see SM3 Table 1. Auditory stimuli were recorded by native German- and French-speakers. Position of the target was randomly assigned among the four possible positions on the screen.

Participants were required to start the test in a quiet room, where they knew they would not be disturbed for the duration of the experiment. The participants started with the consent form, questionnaire and first part of the experiment. Half of the bilinguals had the consent form, questionnaire and started the experiment in German and then had the task in French, the other half started in French. Thanks to a Labvanced feature the screen were calibrated so that each four vertical Arabic number appeared evenly distributed on participant's screen and on 5 X 3 degree of visual angles.

4 Results

4.1 Data analyses and hypothesis testing

All models and a priori hypotheses and general pre-registration can be found here: *https://osf.io/b4p2z/*. The models presented in the results (final models) are obtained by degrading the pre-registered maximal models (Barr et al., 2013). Initial model degradations were done by removing terms with too high correlations in the covariance matrixes (*i.e.* singularity). If the models did not converge, we removed theoretically less significant terms. The formula of all models can be found in the supplementary materials (SM1). Data and model selection procedures can be found at *https://osf.io/b4p2z/.*

Before analyses, we filtered the data on trial level: from initial 57510 trials, 644 trials with RT slower than 3 seconds and or faster than 300 ms were excluded. Trials with RTs +/- 3 standard deviations from each individual mean were also removed (total of 598 trials). In sum, we excluded .02 % of the initial trials before the analyses. After removing error trials (1.15 %), reaction times were analyzed using linear mixed models in RStudio (RStudio Team, 2020) using the packages afex (Singmann, Bolker, Westfall, Aust, & Ben-Shachar, 2022) that integrate the popular lme4 package (Bates et al., 2015). Note that differently to lmer package, within afex models contrasts defaults are sum-to-zero (difference from the grand mean). When necessary, P-values are Bonferroni corrected, all degrees of freedom are estimated by Satterthwaite method. All the following linear mixed models were replicated with log transformed reaction times, resulting in the same significancy patterns. For the ease of interpretation, we report here the analyses on non-transformed reaction times analyses. Errors were not analyzed statistically given they were too few (1.15 %) and that only 92 participants did more than 1 error in the 3 conditions (see SM2). A descriptive table of the type of errors (i.e. which distractors was clicked upon) can be found in the supplementary materials (SM2 Table 3).

We did four main groups of analyses: the first analysis comparing monolinguals, then two analyses comparing bilinguals to language-matched monolinguals (e.g. bilinguals in

German to German monolinguals), and finally we compared the bilinguals' performances in both languages, i.e. German and French. The simultaneous condition was analyzed in a separate model than the Ten- and Unit-first sequential conditions.

For hypothesis testing we integrated the results from the linear mixed models to confirm, reject or adjust the hypothesis H(A) to H(F). With those adjustments we constructed a model to predict RT based on monolingual French performance. This model was then applied to the other two groups and languages for hypothesis testing. For example, from Table 2, the average **1067 ms** for monolingual French in the simultaneous conditions already corresponds to the **intercept** in our model F(A).

Table 2: Reaction times (in ms) for the three conditions for the bilinguals (in German and French) and for German and French monolinguals

		Simultaneous	Ten-first	Unit-first
Group	Language	M(*SD*)		
Monolingual	French	**1,067(*384*)**	877(*364*)	953(*382*)
Monolingual	German	1,077(*411*)	912(*412*)	970(*405*)
Bilingual	German	1,161(*434*)	969(*423*)	1,041(*422*)
	French	1,266(*460)*	1,023(*447*)	1,169(*468*)

Note: standard deviations in parenthesis. Note 1067 corresponds hence to the intercept in F(A)

4.2 Monolingual German vs monolingual French

4.2.1 Simultaneous

The comparison of monolingual French and German groups in the simultaneous conditions was not significant ($F(1,111.53) = .09$, *n.s.*).

4.2.2 Sequential

The model including only the sequential conditions resulted only in an effect of condition ($F(1,106.94) = 37.46$, $p < .001$), indicating the Ten-first was solved faster than the Unit- first by both monolingual groups. Nor the effect of group nor the interactions were significant ($F < .71$).

4.2.3 Hypothesis testing

No significant difference was found between monolinguals' simultaneous nor sequential conditions, hence refuting H(B) about inversion that would have predicted a slower response for monolingual German. The effect of condition and absence of interactions, means both conditions had a similar effect on both groups. In other words, H(C) and H(D) are the same for both monolinguals.

To calculate the adjusted size effect of H(C) and H(D) we subtracted the Ten-first from the Unit-first condition for monolingual French: **877** - **953** ms leading to $\Delta =$ -76, see Table 3. This difference (Δ) could result from three scenarios (1) -76 ms facilitation for the Ten-first, (2) +76 ms interference from the Unit-first or (3) a mix of Ten-first facilitation and Unit-first interference resulting in a total difference of -76 ms between both conditions. We arbitrarily implemented scenario 3 with a -38 ms Ten-first facilitation (leading to 1067 – 38 = ***1029*** ms in the Ten-first condition) and a +38 ms Unit-first interference (leading to 1067 + 38 = ***1105*** ms in the Unit-first condition) , see Table 3. Importantly this arbitrary decision does not affect the predictions, as will be demonstrated in the next step.

Table 3: Model for: $RT = Intercept + \Delta_{\text{Ten} - \text{Unit first}}$

		Simultaneous			Ten-first			Unit-first		
		Meas.	Model		Meas.	Model		Meas.	Model	
Group	Lang.		Pred.	Error		Pred.	Error		Pred.	Error
Mono.	Fr	1067	1067	0	**877**	***1029***	-152	**953**	***1105***	-152
Mono.	Ge	1077	1067	10	912	1029	-117	970	1105	-135

Notes: Model terms are calculated using the formula in the title. Intercept = 1067, Ten- first condition =- -38 ms, Unit- First = +38 ms. Colour code is a continuous scale ranging from red -150, green 0, to red for 150. This colour scale is also used in the following tables. Mono. = Monolinguals, Fr = French, GE = German, Meas. = Measured, Pred. = Prediction. Error = Measure -Model. Sum of absolute errors = 556 ms.

At this point, we can also adjust the weight for the Prime term in F(A). Since we adjust our model using monolingual French as reference, we set the Prime's term from the error of the Ten-First condition (i.e., Model prediction - Measure): ***1029* - 877** = +152 ms**,** see Table 3 and 4.

Table 4: Model for: $RT = Intercept + \Delta_{\text{Ten}-\text{Unit first}}$

		Simultaneous			Ten-first			Unit-first		
		Meas.	Model		Meas.	Model		Meas.	Model	
Group	Lang.		Pred.	Error		Pred.	Error		Pred.	Error
Mono.	Fr	1067	1067	0	877	877	0	953	953	0
Mono.	Ge	1077	1067	10	912	877	35	970	953	17

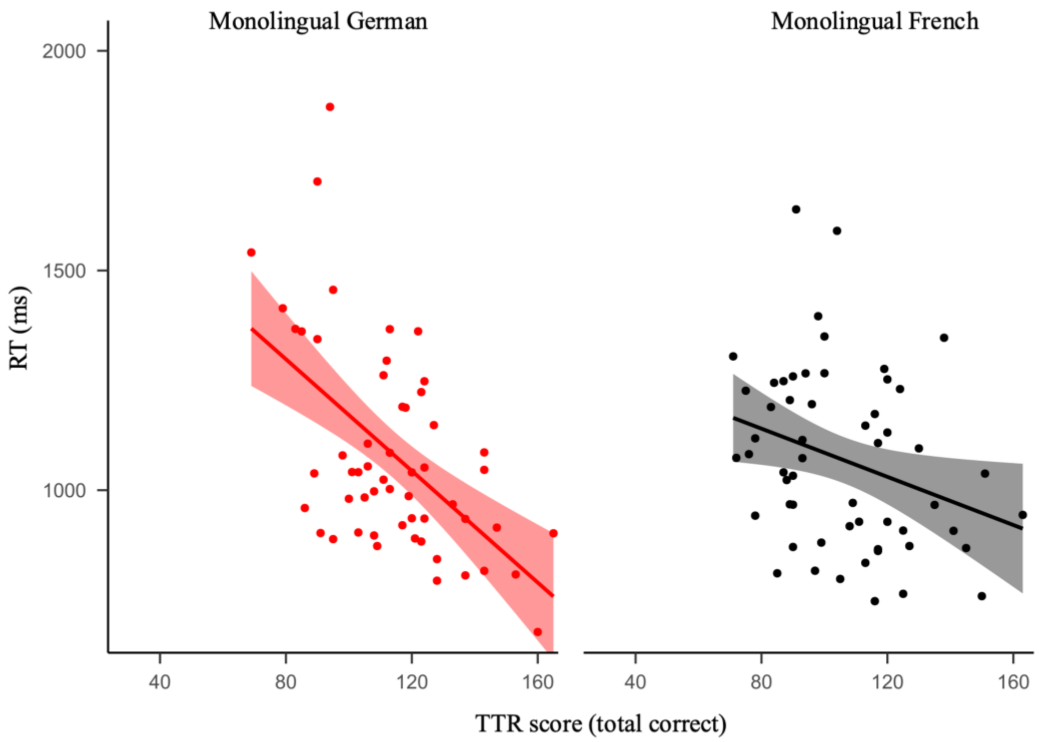
Notes: Prime = 152 ms. Model's Error are calculated by Measure - Model. Sum of absolute errors = 62 ms.

Since the Ten-first, Unit-First and Prime's adjustment are designed to have monolingual French as baseline (i.e. Monolingual French's error = 0) and we make the same predictions for Monolingual German, all the 3 scenarios described above lead to the same predictions for both groups. Hence the arbitrary choice we made (cf. scenario 3 above) is not decisive for later descriptions of this model (for a demonstration see SM5).

4.2.4 Correlations with arithmetic for monolinguals

Pearson correlations between individual mean RT of the simultaneous conditions and TTR resulted in significant negative correlation for monolingual German (r = -0.55 , $t(53)$ = -4.84, p <.001) and monolingual French (r = -0.29 , $t(53)$ = -0.29, p <.05), meaning that the faster simultaneous condition was responded, the better success at the arithmetic test, see Figure 3.

Figure 3: Scatter plot of monolingual German and monolingual French RT and TTR correlations



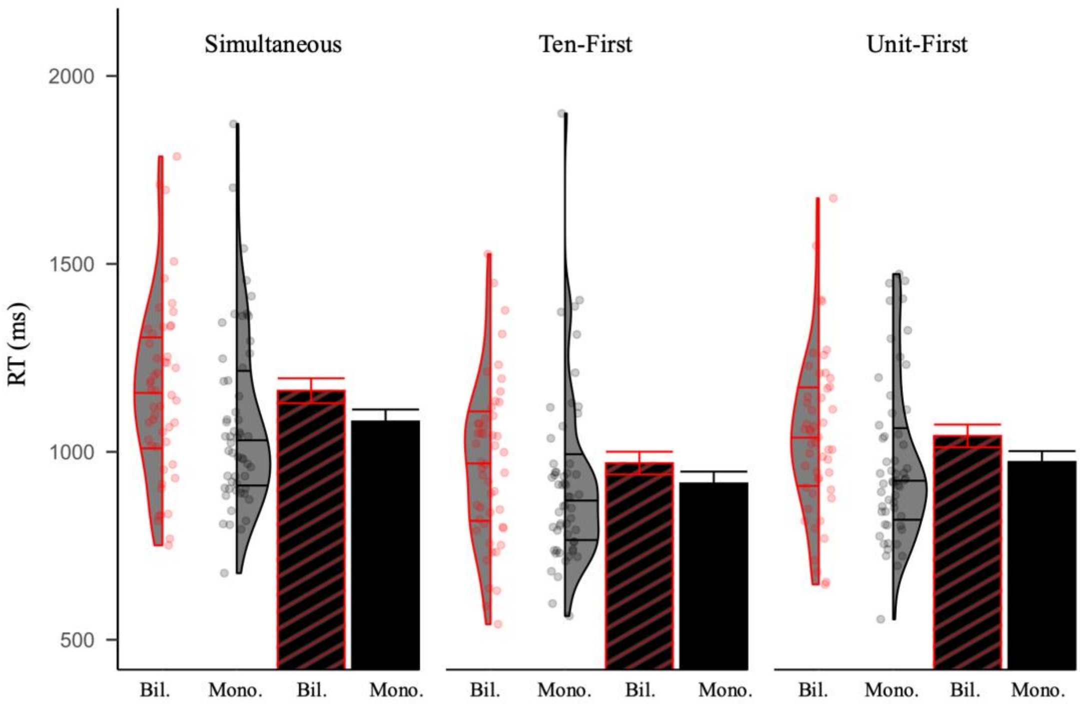
Note: Left panel (red) for German (DE) monolingual, right panel (black) for French (FR) monolingual. Line fitted with a linear regression. TTR is the sum of correct responses. Each point represents one participant.

4.3 Task in German - bilinguals in German vs German monolingual performance

4.3.1 Simultaneous

In the simultaneous condition, bilinguals who did the task in German did not clearly differ from German monolinguals ($F(1, 103.41) = 3.15$, $p = .079$), see Figure 4. Since analyses on log(RT) led to close to significant p-values ($F(1, 102.84) = 3.50$, $p = .06$), we conducted additional Bayesian linear mixed models on this specific comparison with the *brms* package (Bürkner, 2017). Bayesian analyses indicate a $BF_0 = 529$ in favour of the first model compared to the model without the Group effect, meaning the model with groups is 529 times more likely. In other words, bilinguals in German tended to be slower than monolinguals although this difference did not reach significance in frequentist statistical tests.

Figure 4: Bilingual in German vs. German monolinguals



Notes: Each pane represents one of the three conditions. Violinplots with red lines and points correspond to Bilinguals in German, with black lines to German monolinguals. Each point is a participant's average. Striped bar plots represent bilinguals in German, full black bars to German monolinguals, error bars are one standard error. Bil. = Bilinguals, Mono. = Monolinguals.

4.3.2 Sequential

In the sequential conditions, bilinguals in German and German monolinguals did not differ (F (1, 102.95) = 2.29, *n.s.*), nor did group interact with the condition (F (1, 101.42) = 0.42, *n.s.*). Both groups were significantly slower in the Unit-first than the Ten-first condition (F (1.101.42) = 23.87, $p < .001$).

4.3.3 Hypothesis testing

From the results of the simultaneous condition, we confirmed H(E) of a bilingual lexical cost. We obtain the adjustment of H(E), the *bilingual lexical cost*, by subtracting the German simultaneous bilingual to the monolingual condition, **1161-1077** = +84, see Table 5. Hence adjusting our pre-registered value from 60 to 84 ms.

Table 5: Model for: $RT = Intercept + \text{Prime} + \Delta_{\text{Ten} - \text{Unit first}} + Bilingual\ lexical\ cost$

		Simultaneous			Ten-first			Unit-first		
		Meas.	Model		Meas.	Model		Meas.	Model	
Group	Lang.		Pred.	Error		Pred.	Error		Pred.	Error
Mono.	Fr	1067	1067	0	877	877	0	953	953	0
Mono.	Ge	**1077**	1067	10	912	877	35	970	953	17
Bil.	Ge	**1161**	1151	10	996	961	8	1041	1037	4

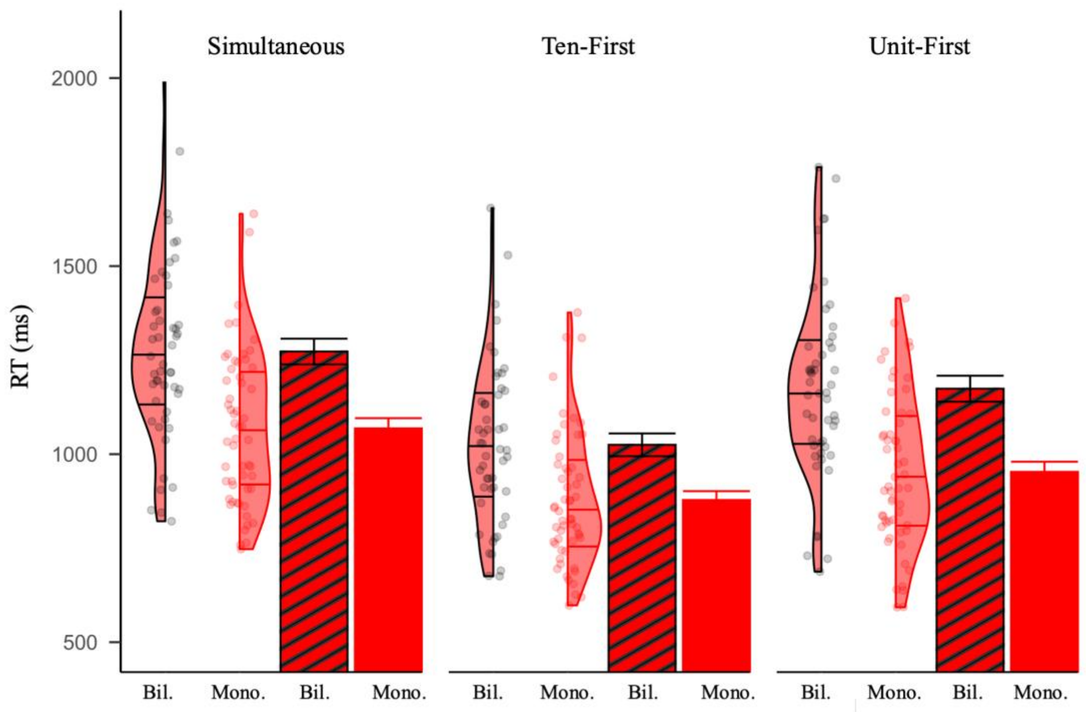
Notes: Bil. = bilingual. Bilingual lexical cost = 84 ms for all bilinguals. Model's Error are calculated by Measure - Model. Sum of absolute errors = 84 ms.

4.4 Task in French - French bilinguals vs French monolinguals performances

4.4.1 Simultaneous

In French bilinguals were on average 199 ms slower than French monolinguals in the simultaneous condition ($F(1,110.57) = 21.54$, $p < .001$), see Figure 5.

Figure 5: Bilingual in French vs. French monolinguals



Notes: Each pane represents one of the three conditions. Violinplots with black lines and points correspond to Bilinguals in German, with red lines to German monolinguals. Each point is a participant's average. Striped bar plots represent bilinguals in French, full black bars to French monolinguals, error bars are one standard error. Bil. = Bilinguals, Mono. = Monolinguals.

4.4.2 Sequential conditions

We found a main effect of group in the sequential conditions (F (1,108.38) = 20.92. p <.001): bilingual in French were slower than monolingual French, hence further confirming the results from the simultaneous condition. A main effect of condition was also found (F(1,103.95) = 152.78, p <.001) and critically a significant interaction between Group and Condition (F(1,103.95) = 17.57, p <. 001).

Post-hoc of the estimated marginal means of this interaction indicate that bilinguals were slower than monolinguals in both sequential condition: 146 ms for the Ten-first (t(108.71) = 3.81, p = .001), and 216 ms for the Unit-first (t(107.71) = 5.06, p < .001). Also, the Ten-first condition was solved faster than the Unit-first condition by both groups: 76 ms faster for Monolingual French, ($t(102.97)$ = - 6.17, p < .001) and 146 ms for bilinguals (t(104.83) = -11.55, p < .001)). However, these post-hoc tests are exclusively informative about significant differences between conditions, but not about effect size differences of these conditions hence complicating hypothesis testing which contained predictions concerning not only the direction but also the amplitude of the expected effects.

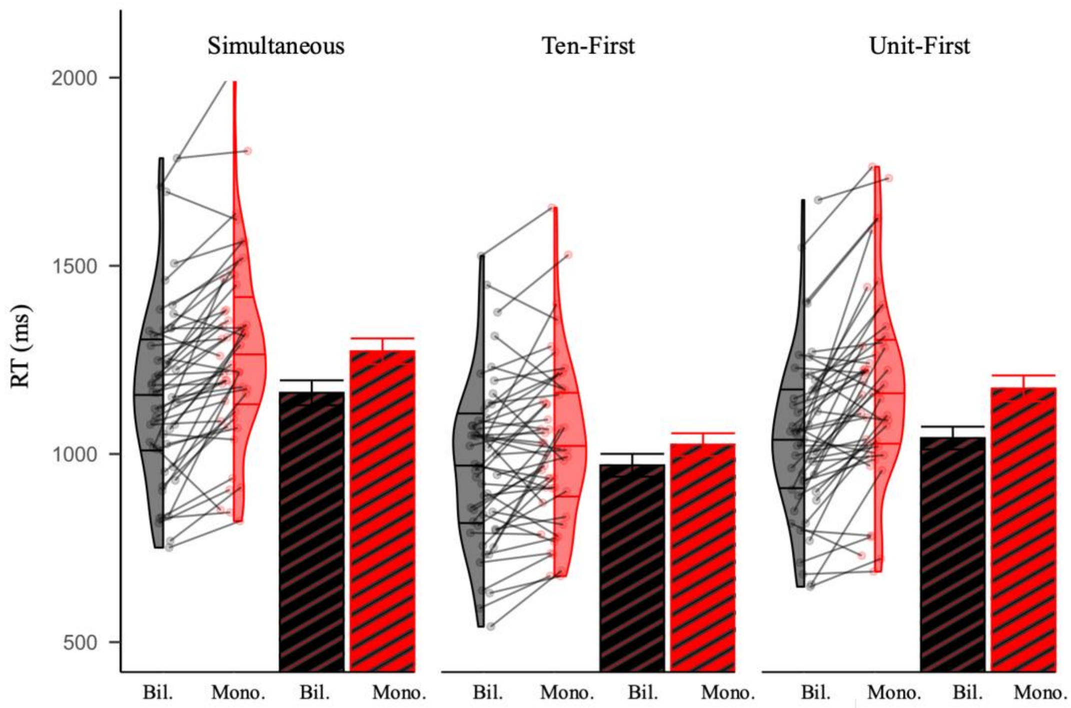
4.5 Bilinguals performing in German vs French

To compare bilingual's performances in German and French full within-subject analyses were conducted.

4.5.1 Simultaneous

The simultaneous condition in bilinguals resulted in an effect of language: (F (1,47.45) = 37.44. p <.001), as bilinguals were 105 ms faster in German (LM1) than in French (LM2).

Figure 6 : Bilingual's Reaction in each language for each condition



Notes. Each pane represents one of the three conditions. Violinplots with black lines and points correspond to Bilinguals in German, with red lines to Bilinguals in French. Each point is a participant's average, lines connect each participant within conditions. Black barplots with red lines correspond to bilinguals in German, Red filled barplots with black stripes to French bilinguals, error bars are one standard error. Bil.GE = Bilinguals in German, Bil.FR = Bilinguals in French.

4.5.2 Sequential

The sequential condition confirmed the effect of language (F (1,48.94) = 33.82. p <.001). Like for the previous analyses we found an effect of condition in bilinguals (F (1,50.59) = 83.55, p <.001). Group and condition significantly interacted (F (1,48.78) = 12.91. p <.001)) and we conducted post-hoc paired tests on the estimated marginal means.

Post-hocs showed that bilinguals were faster in German than in French for both conditions: 54 ms faster in the Ten-first condition (t(48.94) = -3.21, p <.05) and 128 ms in the Unit-first (t(48.85)) = -6.17, p <.001)). Furthermore, bilinguals solved the Ten-first condition faster than the Unit-first condition in both languages: in German they were 72 ms faster (t(58.09) = -4.38, p <.001) in French they were 146 ms faster (t(58.14) = -9.45, p < .001).

4.5.3 Hypothesis testing

Comparing bilinguals' performances in German and French performance directly tested the *LM2 cost* hypotheses (i.e. H(F)). By subtracting the simultaneous condition in French and in German, **1266 -1161** = 105 ms (see Table 6), we obtain the weight of the *LM2 cost*. Hence from the initial model F(A), by removing non-significant terms through the analyses, we obtained the following model with the previous terms, see F(B) in Table 6.

Table 6: Model F(B): $RT = Intercept + Prime + \Delta_{\text{Ten}-\text{Unit first}} + Bilingual\ lexical\ cost + Bil\ LM2\ cost$

		Simultaneous			Ten-first			Unit-first		
		Meas.	Model		Meas.	Model		Meas.	Model	
Group	Lang.		Pred.	Error		Pred.	Error		Pred.	Error
Mono.	Fr	1067	1067	0	877	877	0	953	953	0
Mono.	Ge	1077	1067	10	912	877	35	970	953	17
Bil.	Ge	**1161**	1151	10	996	961	8	1041	1037	4
Bil.	Fr	**1266**	1256	10	1023	1066	**-43**	1169	1142	**27**

Notes: Adjusted the Bil LM2 cost = 105 ms for French bilinguals. Model's Error are calculated by Measure- Model. Sum of absolute errors = 164 ms.

F(B) however fails to consider the significant interactions between group and condition found in the comparison of bilinguals in French with monolingual French and the comparison of languages in bilinguals we described above. Therefore, we added some terms to the theoretical model (F(A)). This *a posteriori* terms are for the Ten- and Unit-first conditions of bilinguals in French, see F(C).

$$F(C): RT = RT$$
$$= Intercept + Prime + \Delta_{Ten-Unit\ first} + Bilingual\ lexical\ cost$$
$$+ Bilinguals\ LM2\ cost + Bilingual\ LM2|French|Ten\ first$$
$$+ Bilingual\ LM2|French|Unit\ first$$

The weight of the new terms of (F(C)) are deduced from Table 7: a **-43** ms facilitation for $Bilingual\ LM2|French|Ten\ first\ cost$ and a **27** ms interference for $Bilingual\ LM2|French|unit\ first$. Applying these terms to F(C), we obtain the model in Table 7.

Table 7: Model for: $Intercept + Prime + \Delta_{Ten\,-Unit\,first} + Bilingual\ lexical\ cost + Bilinguals\ LM2\ cost + Bilingual\ LM2|French|Ten\ first\ cost + Bilingual\ LM2|French|Unit\ first$

		Simultaneous			Ten-first			Unit-first		
		Meas.	Model		Meas.	Model		Meas.	Model	
Group	Lang.		Pred.	Error		Pred.	Error		Pred.	Error
Mono.	Fr	1067	1067	0	877	877	0	953	953	0
Mono.	Ge	1077	1067	10	912	877	35	970	953	17
Bil.	Ge	**1161**	1151	10	996	961	8	1041	1037	4
Bil.	Fr	**1266**	1256	10	1023	1023	0	1169	1169	0

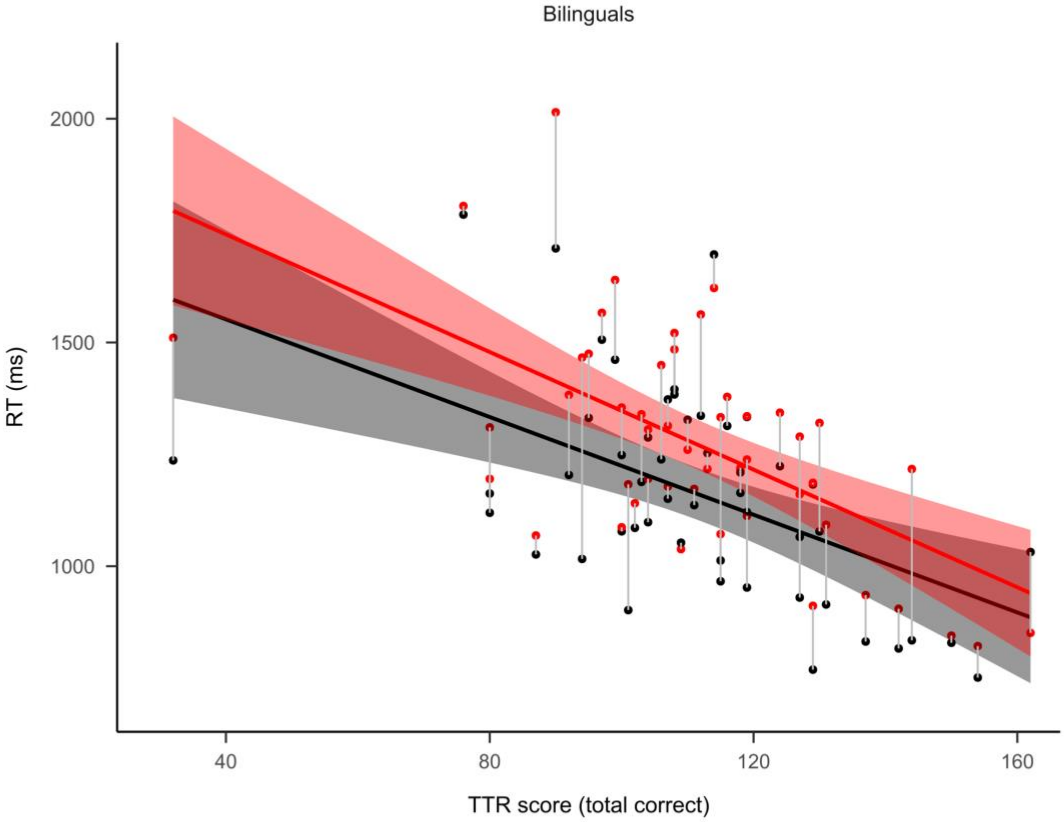
Notes: Added a specific -43 ms facilitation for French bilinguals in the Ten-first condition and 27 interferences for Unit-First. Sum of absolute errors = 94 ms.

In conclusion, contrary to our predictions, bilingual's languages interacted with the conditions, such that bilinguals in French were faster than initially predicted in the Ten-first condition and slower than initially predicted in the Unit-first condition.

4.5.4 Correlation with arithmetic for bilinguals

We conducted Pearson correlations between individual mean RT of the simultaneous conditions in German and in French with the TTR. Bilinguals' TTR scores correlated with RT in the simultaneous condition in German (r = -0.51, $t(48)$ = -4.11, p <.001) and in French (r = -0.59, $t(48)$ = -5.10, p <.001), see Figure 6.

Figure 6: Bilinguals correlation between each language's simultaneous condition and TTR



Note 1: Black line = regression for responses in German and TTR, red in French) Grey lines connected points corresponding to single participants.

5 Discussion

We compared German-French bilinguals to German and French monolinguals in an auditory-visual number matching task. In this task, participants heard a two-digit number (*i.e.* /forty-two/) followed by four visually presented two-digit Arabic numbers: one target and three distractors (*i.e.* 42, 24, 43 and 34). There were three different priming conditions, one simultaneous and two sequential ones. In the simultaneous condition participants had to match the heard numbers with non-primed (i.e., simultaneously) presented two-digit Arabic numbers (*i.e.,* 42, 24, 43 and 34). In the two sequential conditions one part of the two-digit number information was revealed 500 ms before as a prime: either Ten-first (*i.e.,* 4_, 2_, 4_ and 3_) or Unit-first (*i.e.,* _2, _4, _3 and_4). The results show first that bilinguals in German (LM1) were slightly slower than German monolinguals in all conditions. Second, bilinguals in French (LM2) were slower than French monolinguals, and compared to when they did the same tasks in German. Third, bilinguals in French were relatively faster than expected for the Ten-First condition and slower for the Unit-first condition. In the following we will discuss the theoretical implications and for each group comparison.

5.1 Monolingual French vs German monolingual

The comparison of monolingual German and French revealed that, independently from the group or language, the Ten-first condition was solved faster than the Unit-first condition. A simple cognitive explanation might be provided by visual matching between formats: when the participants hear the number word, the correspondent mental representation of the visual Arabic number is automatically activated (/zwei und vierzig/ or /quarante-deux/ → 42). Since the ten information is also visually the first to be seen in western left-to-right reading, the condition presenting the Ten-first is facilitated. This is likely also due to writing procedures, since the ten is the first number to be written before the unit (see for example Moeller, Shaki, et al., 2015). Given that we did not find any interactions between group and condition, our design did not reveal an effect of morpho-syntactic inversion on verbal-visual matching in adult German monolinguals. These results are in line with Steiner, Banfi, et al., (2021) and (Poncin et al., 2019), suggesting that morpho-syntactic inversion in German might affect number transcoding in monolingual children but not adults, as tested here. Hence this refutes two specific hypotheses for monolinguals: larger facilitation of the Unit-first in German H(D) and larger facilitation for the Ten-first in French H(C). Similarly, the absence of a difference between groups in simultaneous and sequential conditions furthermore rejects our hypothesis about an effect of inversion in German H(B). In general, the absence of an influence of the morpho-syntactic ten-unit inversion in German might show that transcoding in adults is automatized, *i.e.* that monolinguals retrieve number words directly from long-term memory as predicted in the ADAPT model (Barrouillet et al., 2004). Hence, when a number word is heard, the corresponding visual Arabic number is automatically activated, bypassing morpho-syntactic interferences. Another possibility, however, is that the effect of morpho-syntactic inversion in German is too small and/or the task not sensitive enough to detect it. For example, the 500 ms cuing used here might be too long to elicit an observable effect in adults. Indeed, although only descriptively and only on a proportion of individuals, monolingual German confounded the target with the inversion distractor in total 86 times vs. 24 times for monolingual French (see SM 2 Table 3). Hence, the effect of inversion in adult monolinguals might require more fine-tuned experimental designs to become visible (Nuerk et al., 2005).

5.2 Bilinguals *vs* monolinguals in German

How does bilinguals' performances compare to those of language-matched monolinguals? We compared bilinguals in German and German monolinguals on the simultaneous and sequential conditions. The simultaneous condition is an indirect test of the *bilingual lexical cost* H(E) since it compared bilingual's L(M)1 with monolinguals. The hypothesis states that bilinguals are generally slower than monolinguals. Our experiment provides partial support for the *bilingual lexical cost* hypothesis. Indeed, comparing both groups in the simultaneous conditions led to heterogeneous results: linear mixed models resulted in marginally significant differences ($p = .08$ with reaction times and $p = .06$ with logarithm of the reaction times), while Bayesian analyses indicate that a difference between monolinguals and bilinguals is 529 times more likely than no group difference. However, we did not find any main effect of group in the analyses of the sequential conditions. This *bilingual lexical cost* was theoretically predicted as a result of either bilingual's language coactivation needing additional control mechanisms which are not required in monolingual processing (Green, 1998) or decreased absolute frequency of use given that L1 use is mutually exclusive of L2 use, hence reducing frequency compared to undivided L1 use in monolinguals (Dijkstra & Heuven, 2002). We might have missed a clearly significant *bilingual lexical cost* due to a lack of measure sensitivity. For example, we used a recognition task, while production tasks such as number naming are known to be more sensitive to language contrasts within bilinguals (Vander Beken & Brysbaert, 2018). Overall, the inconclusive results of a difference between bilinguals and monolinguals in the simultaneous condition and the similar pattern in the sequential conditions are in line with the high proficiency level in German of the present bilingual sample (probably partially explained by the linguistic proximity between German (LM1) and Luxembourgish (native L1). Furthermore, both groups appear to be affected similarly by the sequential conditions as they are solving the Ten-first condition faster than the Unit-first condition. This furthermore suggests that our bilingual sample in German does not appear to benefit from knowing French, where number words start with tens. In other words, there does not seem to be a transfer from the LM2 to the LM1 morpho-syntax. A recent study reported morpho-syntactic effects from the L2 affecting L1 and some of the effects were observable on accuracy rather than reaction time's, Furthermore, artificially constructed number words were used, which could induce stronger morpho-syntactic interferences than with the present design (Xenidou-Dervou et al., 2023), explaining why such effects were not observed here.

5.3 Bilinguals' *vs* monolingual in French

Bilinguals were clearly slower than monolinguals in French. This difference was hypothesized by the cumulation of a *bilingual lexical cost* H(E) discussed above and *LM2 cost* H(F) for bilinguals in French, as their LM2. The *LM2 cost* replicates previous studies on different samples with this specific language profile using various tasks: simple and complex arithmetic (Van Rinsveld et al., 2015), two-digit number matching (Lachelin et al., 2022) and single digit number reading (Lachelin et al., 2023), see also (Garcia et al., 2021) for a meta-analysis. The LM2 cost indicates that the second language of learning mathematics requires longer time to activate, this despite French being an important medium for schooling and being used as language of math instruction for 7 years (vs 6 years of math instruction in German). The LM2 cost generally aligns with results from language switching costs (i.e. Bernardo, 2001; Grabner et al., 2012), given that number words are likely learned in German. Hence it possibly stems from slower lexical access in the L(M)2 compared to the L(M)1, as predicted by several bilingual models (i.e. Dijkstra & Heuven, 2002) and specifically for bilingual numbers (Lachelin et al., 2023). The comparison of the sequential conditions in French revealed an interaction between bilinguals and monolinguals conditions, which was not found in German. This interaction suggests that bilinguals in French, other than in German, responded differently than French monolinguals to the sequential conditions. Given the additivity between the LM2 cost and a possible effect of conditions (see F(A)), these effects could only be disentangled by considering additional findings from other condition comparisons as will be discussed in § 4.5 Prediction model F(C).

5.4 Bilinguals

Within-participant comparisons showed that bilinguals were significantly faster in German than in French (105 ms) for the simultaneous condition, which is the most direct measure and confirmation for the *LM2 cost* H(F). Critically, as for the comparison between French monolinguals and bilinguals in French, we found an interaction between sequential conditions and language. However, these statistics are not informative about differences in effect sizes which could arise from the addition of different hypothesized effects, see F(A). To test our hypothesis, we proceeded by dropping the terms of F(A) that were not significant and adjusting the significant hypothesized terms with the obtained measures. Conceptually, we built a model based on monolingual French and applied it to other groups and adjusted significant terms *a posteriori* based on the measures we obtained. After deleting non-significant terms, we ended with an adjusted model, F(B).

5.5 Conclusions from building model F(C) *a posteriori*

Model F(B), using monolingual French performances as baseline can predict the results for bilinguals in German, suggesting that both monolingual groups responded similarly to the conditions. However, when applied to bilinguals in French, the model remains with an error between the model and the measures of -43 ms for the Ten-first and 27 for the Unit-first condition, see Table 6. Hence, we have built a theoretical *a posteriori* model F(C) which can predict the data parsimoniously with seven terms and leaving a total of 94 ms total unexplained error between the measures and the model, see SM6. F(C), predicts that bilinguals in French have a relative facilitation for the Ten-first and an interference for the Unit-first condition.

This model is justified by the significant interactions in the between-subject comparison of bilinguals to monolinguals in French and the within-subject comparison of bilinguals in French vs German. The theoretical implication of model F(C) is that bilinguals' task performance in French (LM2) was differently influenced by the (sequential) conditions than French monolinguals, but this was not the case when bilinguals performed the task in German (LM1). This could be because later verbal representations of number might be more malleable than earlier acquired ones.

In LM2, the Ten-first and Unit-first conditions revealed enhanced relative facilitation and interference effects. On one hand, the Ten-first increased relative facilitation in LM2 suggests that bilinguals can flexibly take advantage from the LM2's specific linguistic cues. We interpret this advantage as originating from the LM2 morpho-syntax transparency relative to the visual place value order of Arabic numbers in French (i.e. *Ten*-unit as in *4*2). On the other hand, the condition where we cued the unit (i.e. Unit-first) following LM1's inverted morpho-syntax seemed to cause additional interference when processing numbers in the LM2. A possible explanation for this is an over-generalization or automatization of identification of the first part of a number word as a unit (as done in German).

5.6 Summary

The presence of two languages in sequential bilinguals results in an LM2 cost in comparison to the LM1. This LM2 cost is added on top of an LM1 bilingual lexical cost when compared to monolinguals. However, bilinguals appear to flexibly adapt number processing across languages that are inverted or not respective to Arabic numbers. Since processing in the LM1 resembled that of monolinguals, we found this impact from the LM1onto the LM2 but not from the LM2 onto the LM1. On one hand the ten-unit morpho-syntactic inversion of the LM1's seems to permeate into LM2 processing. Such that after hearing a two-digit number in the LM2 and visually cuing the unit part of the number interferes with transcoding processes in LM2. When bilinguals hear a two-digit number in LM2 they might thus co-activate the verbal (inverted) form in LM1 which then interferes with the response. On the other hand, the ten-unit morpho-syntactic congruence of the LM2 with the visual arabic number place value system elicits a relative advantage in bilinguals. In other terms, linguistic cues from the LM2 that are coherent with visual input's morpho-syntax (i.e. in ten-unit order) can facilitate processing in the LM2. Note that LM2 co-activation does not seem to happen when hearing LM1 number-words, otherwise we should have found a facilitation for the Ten-first condition in German.

All three groups arithmetic's TTR score correlated with reaction times of the simultaneous condition for both monolingual and for bilinguals in German and in French. This correlation indicates that the faster participants matched the numbers, the more arithmetic problems they were able to solve. This result confirms previous results suggesting common processes involved for both arithmetic and transcoding (Steiner, Banfi, et al., 2021; Xenidou-Dervou et al., 2015). It indicates that the above-described cognitive mechanisms in bilinguals might also be at hand for solving arithmetic.

5.7 Limitations

Language profiles were not perfectly homogeneous within the groups and the bilingual group was rather multilingual with an average 4.64 languages (see Table 2). However, one of the additional languages is Luxembourgish, which shares phonological and morpho-syntactic similarities with German. Most of the bilingual group also indicated English as a fourth language. Furthermore, the monolingual groups could also be considered as late unbalanced bilinguals, given that all reported speaking at least 2 languages, mostly English which is learned at school and exposed through internet. Another important aspect characterizing the language profiles of our bilingual sample is that they are sequential learners, hence our results might not generalize to early simultaneous bilinguals. Nevertheless, this sample is homogeneous regarding the formal acquisition of the L(M)2 by learning all school subjects in that language at the same age, a situation closely resembling language immersion.

5.8 Conclusion

The present results indicate that the morpho-syntactic rules have larger effects on a more malleable LM2, rather than the LM1. The present results underline the importance of developing cognitive models that include bilingualism, since bilinguals tend to perform tasks differently than monolinguals in both of their languages and bilingual's L(M)2 shows complex interactions with the L(M)1 (Lachelin et al., 2023). In conclusion, the present study provides new insights into how sequential bilinguals compare to monolinguals with regards to two-digit number-word processing when the languages differ in the morpho-syntactic ten/unit place-value position relative to Arabic-numbers (i.e., inversion). Bilinguals do not only tend to be slower in lexical access for both of their languages compared to monolinguals, but they display unique adaptation to the morpho-syntactic specificities of their LM2 regarding number words.

6 Supplementary

6.1 Supplementary material 1

6.1.1 Model/Formula:

$$F(A): RT$$
$$= Intercept + Prime + Inversion + Ten\ first \mid Language + Unit\ first|Language$$
$$+ Bilingual\ lexical\ cost + Bilinguals\ LM2\ cost$$

To have a sizeable table, we have abbreviated the formula's term as follows:

$$F(A): RT = Ipt + P + Inv + Tf|\ Lang + Uf|Language + BLc + BLM2c$$

6.1.2 Hypothesized Effect size:

Term	Effect size:
Ipt	900
P	0
Inv	30
Tf|FR	-100
Tf|DE	-30
Uf|FR	30
Uf|DE	-60
BLc	60
BLMc	60

Notes: effect sizes in ms

6.1.3 Prediction:

SM1 Table 1: Formula and numerical effect size prediction for each cell:

Group	Language	Conditions					
		Simultaneous		Ten first		Unit First	
MonoFR	FR	Ipt	900	Ipt + P + Tf|FR	740	Ipt + P + Uf|FR	870
MonoDE	DE	Ipt + Inv	930	Ipt + P + Inv + Tf|DE	840	Ipt + P + Inv + Uf|DE	810
Bilinguals	DE	Ipt + Inv + BLc	990	Ipt + P + Inv + Tf|DE + BLc	900	Ipt + P + Inv + Uf|FR + BLc	870
Bilinguals	FR	Ipt + Inv + BLc + BLM2c	1020	Ipt + P + Inv + Tf|FR + BLc + BLM2c	860	Ipt + P + Inv + Uf|FR+ BLc + BLM2c	990

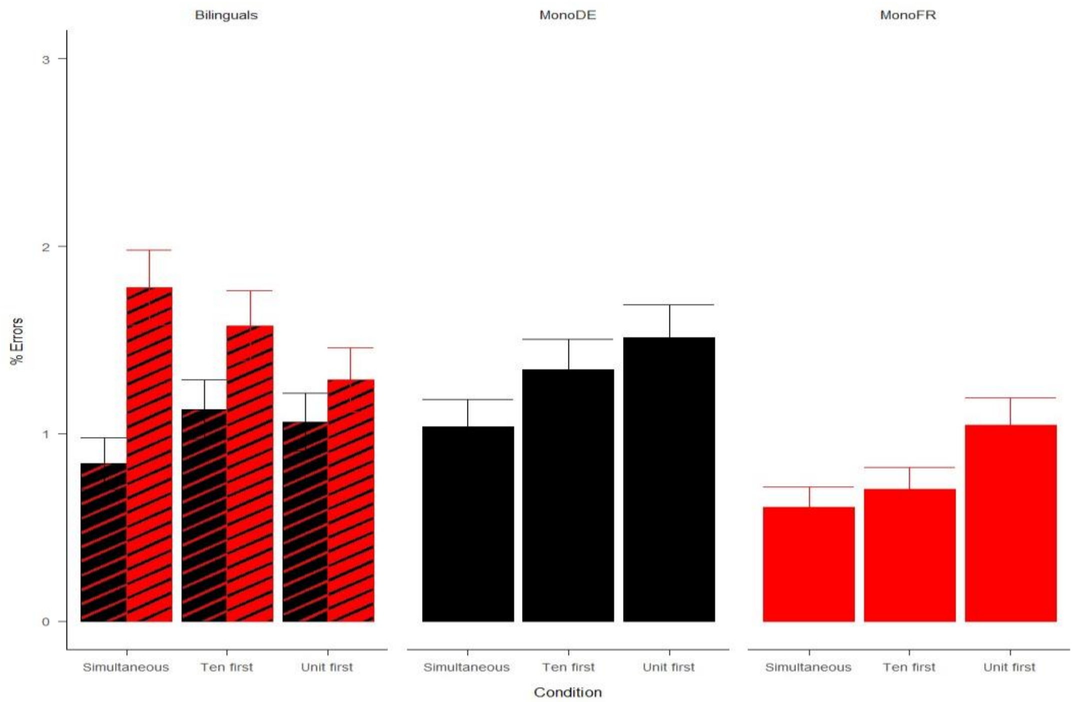
Notes: FR = French, DE = German. Each condition has on one size the part of the formula that predicts it and the prediction.

6.2 Supplementary material 2

6.2.1 Error analyses

We found a total 1.15% error trials, reflecting the relative simplicity of the task and attention of the participants. 32 participants did 0 errors and 37, 1 error over the 3 conditions.

SM1 Figure 1: % Of errors for each group across languages and conditions.



Notes: Black = German, Red = French. Stripped bars are bilinguals, those filled in black are the conditions in German. Filled in red conditions in French.

SM1 Table 1: Average % errors

Group	Language	Simultaneous		Ten first		Unit first	
		M	sd	M	sd	M	sd
Monolingual	French	.61	7.77	.70	8.35	1.05	10.17
Monolingual	German	1.04	10.13	1.34	11.50	1.51	12.20
Bilingual	German	.84	9.13	1.13	10.56	1.06	10.25
Bilingual	French	1.78	13.22	1.57	12.45	1.29	11.28

Note: M = mean, sd = standard deviation

SM1 Table 2: Number of errors per conditions (absolutes)

Group	Language	Conditions		
		Simultaneous	Ten first	Unit first
Monolingual	French	30	35	52
Monolingual	German	50	65	73
Bilingual	German	37	50	47
Bilingual	French	77	69	56

Note: Total (absolute) number of errors per conditions

SM1 Table 3: Types of errors

Group	Language	Type of errors		
		Inversion distractor	Ten distractors	Unit distractors
Monolingual	French	24	68	25
Monolingual	German	86	73	29
Bilingual	German	64	51	19
Bilingual	French	60	48	94

Note: Total (absolute) number of errors, per error types (i.e. type of distractors clicked).

6.3 Supplementary material 3

SM2 Table 1: Stimuli's list

Target	Inversion Distractor	Ten Distractor	Unit Distractor
23	32	24	13
24	42	25	34
25	52	26	35
26	62	27	36
27	72	28	37
28	82	29	38
32	23	31	42
34	43	35	24
35	53	36	45
36	63	37	46
37	73	38	47
38	83	39	48
42	24	43	52
43	34	42	53
45	54	46	35
46	64	47	56
47	74	48	57
48	84	49	58
52	25	53	62
53	35	54	63
54	45	53	64
56	65	57	46
57	75	58	67
58	85	59	68
62	26	63	72
63	36	64	73
64	46	65	74
65	56	64	75
67	76	68	57
68	86	69	78

Notes: list of all stimuli used. The second half of the stimuli had ten distractors being unit distractors -1 and ten distractors -10.

6.4 Supplementary material 4

6.4.1 Monolingual German vs monolingual French

6.4.1.1 Simultaneous

SM Table 7: : RT ~ 1 + Group + (Group | Stim) + (1 | ID)

Effect	df	F	P values
Group	1, 111.53	.09	.77

Note: Here the Group factor corresponds to the Language factor: monolingual German and monolingual French.

6.4.1.2 Sequential

SM Table 8: RT ~ 1 + Group * Condition + (Group | Stim) + (Condition | ID)

Effect	df	F	P values
Group	1, 110.64	.59	.44
Condition	1, 106.94	37.46	< .001
Group:Condition	1, 106.94	.71	.40

6.4.2 In German - bilinguals in German vs German monolingual performance

6.4.2.1 Simultaneous

Table 4: Bilinguals in German and German monolinguals, simultaneous condition

Effect	df	F	P value
Group	1, 103.41	3.15	.079

Note. Final model's formula: 1 + Group + (1 | Stim)

6.4.2.2 Sequential

Table 5: Bilinguals in German and German monolinguals, sequential conditions

Effect	df	F	P value
Condition	1, 95.90	23.87 ***	<.001
Group	1, 102.95	2.29	.134
Condition:Group	1, 101.42	0.41	.521

Note. Final model's formula: 1 + Condition * Group + (Condition | Stim) + (Condition | ID)

6.4.3 In French - French bilinguals vs French monolinguals performances

6.4.3.1 Simultaneous

Table 6: Bilinguals in French and French monolinguals, simultaneous condition

Effect	df	F	P value
Group	1, 110.57	21.54 ***	<.001

Note. Final model's formula: 1 + Group + (Group | Stim) + (1 | ID)

6.4.3.2 Sequential conditions

Table 7 : Bilinguals in French and French monolinguals, sequential conditions

Effect	df	F	P value
Condition	1, 103.95	152.78 ***	<.001
Group	1, 108.38	20.92 ***	<.001
Condition:Group	1, 103.95	17.57 ***	<.001

Notes: final model's formula: 1 + Condition * Group + (Group | Stim) + (Condition | ID)

6.4.4 In bilinguals

6.4.4.1 Simultaneous

Table 8: Bilinguals in German and bilinguals in French: simultaneous condition

Effect	df	F	P value
Language	1, 47.45	37.44 ***	<.001

Notes. Final model's formula: 1 + Language + (1 + Language \| Stim) + (1 + Language \| ID)

6.4.4.2 Sequential

Table 9: Bilinguals in German and bilinguals in French: sequential conditions

Effect	df	F	P value
Condition	1, 50.59	83.55	<.001
Language	1, 48.94	33.82	<.001
Condition:Language	1, 48.78	12.91	<.001

Notes. Final model's formula: 1 + Condition * Language + (Condition | Stim) + (Condition + Language + Condition:Language | ID)

6.5 Supplementary material 4

In the following we demonstrate that independently from the scenario used for $\Delta_{\text{Ten}-\text{Unit first}}$ the model's errors and predictions, i.e. F(B) -measures remain unchanged. The three scenarios are (1) Ten-First = -74, (2) Unit-First = 74, (3) Ten-First = -36 AND Unit-First = 36.

$$RT = Intercept + Prime + \Delta_{\text{Ten}-\text{Unit first}} + Bilingual\ lexical\ cost + Bil\ LM2\ cost$$

6.5.1 Scenario 1

Model (1): $RT = Intercept + Prime + (\boldsymbol{Ten - First}) + Bilingual\ lexical\ cost + Bil\ LM2\ cost$

Group	Language	Simultaneous	Ten-first	Unit-first
MonoFR	French	1067	877	953
MonoDE	DE	1067	877	953
Bilinguals	DE	1151	961	1037
Bilinguals	FR	1256	1066	1142

			Error: Model- Measure				
Terms	Weights:						
$Intercept$	1067		Group	Language	Simultaneous	Ten-first	Unit-first
$Prime$	-114		MonoFR	French	0	0	0
$Ten - First$	-76		MonoDE	DE	10	35	17
$Bilingual\ lexical\ cost$	84		Bilinguals	DE	10	8	4
$Bil\ LM2\ cost$	105		Bilinguals	FR	10	-43	27

TOTAL ERROR: 164

6.5.2 Scenario 2

$$Model\ (2): RT = Intercept + Prime + (\boldsymbol{Unit - First}) + Bilingual\ lexical\ cost + Bil\ LM2\ cost$$

			Group	Language	Simultaneous	Ten-first	Unit-first
			MonoFR	French	1067	877	953
			MonoDE	DE	1067	877	953
			Bilinguals	DE	1151	961	1037
			Bilinguals	FR	1256	1066	1142
			Error: Model- Measure				
Terms	Weights:						
$Intercept$	1067		Group	Language	Simultaneous	Ten-first	Unit first
$Prime$	-114		MonoFR	French	0	0	0
$Unit - First$	76		MonoDE	DE	10	35	17
$Bilingual\ lexical\ cost$	84		Bilinguals	DE	10	8	4
$Bil\ LM2\ cost$	105		Bilinguals	FR	10	-43	27
TOTAL ERROR: 164							

6.5.3 Scenario 3

$$Model(3): RT = Intercept + Prime + \Delta_{\text{Ten}-\text{Unit first}} + Bilingual\ lexical\ cost + Bil\ LM2\ cost$$

			Group	Language	Simultaneous	Ten-first	Unit-first
			MonoFR	French	1067	877	953
			MonoDE	DE	1067	877	953
			Bilinguals	DE	1151	961	1037
			Bilinguals	FR	1256	1066	1142
			Error: Model- Measure				
Terms	Weights:						
$Intercept$	1067		Group	Language	Simultaneous	Ten-first	Unit-first
$Prime$	-114		MonoFR	French	0	0	0
$Ten - First$	-36		MonoDE	DE	10	35	17
$Unit - First$	36		Bilinguals	DE			
$Bilingual\ lexical\ cost$	84		Bilinguals	FR	10	8	4
$Bil\ LM2\ cost$	105		Group	Language	10	-43	27
				TOTAL ERROR: 164			

6.6 Supplementary Material 6

Being a dynamic excel table this material is not suited to be integrated in this thesis. It can nevertheless be found at: *https://osf.io/b4p2z/.*

7 References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*, *128*(3), 466–478. https://doi.org/10.1016/j.actpsy.2008.03.014

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Bahnmueller, J., Moeller, K., Mann, A., & Nuerk, H.-C. (2015). On the limits of language influences on numerical cognition – no inversion effects in three-digit number magnitude processing in adults. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01216

Bahnmueller, J., Nuerk, H.-C., & Moeller, K. (2018). A Taxonomy Proposal for Types of Interactions of Language and Place-Value Processing in Multi-Digit Numbers. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01024

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00328

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004a). ADAPT: A Developmental, Asemantic, and Procedural Model for Transcoding From Verbal to Arabic Numerals. *Psychological Review*, *111*(2), 368–394. https://doi.org/10.1037/0033-295X.111.2.368

Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004b). ADAPT: A Developmental, Asemantic, and Procedural Model for Transcoding From Verbal to Arabic Numerals. *Psychological Review*, *111*(2), 368–394. https://doi.org/10.1037/0033-295X.111.2.368

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*(3), 201–221. https://doi.org/10.1016/S0010-0277(02)00178-6

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bernardo, A. B. I. (2001). Asymmetric activation of number codes in bilinguals: Further evidence for the encoding complex model of number processing. *Memory & Cognition*, *29*(7), 968–976. https://doi.org/10.3758/BF03195759

Bernardo, A. B. I. (2005). Language and Modeling Word Problems in Mathematics Among Bilinguals. *The Journal of Psychology*, *139*(5), 413–425. https://doi.org/10.3200/JRLP.139.5.413-425

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General*, *124*(4), 434–452. https://doi.org/10.1037/0096-3445.124.4.434

Bürkner, P.-C. (2017). **brms**: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Bylund, E., Antfolk, J., Abrahamsson, N., Olstad, A. M. H., Norrman, G., & Lehtonen, M. (2023). Does bilingualism come with linguistic costs? A meta-analytic review of the bilingual lexical deficit. *Psychonomic Bulletin & Review*, *30*(3), 897–913. https://doi.org/10.3758/s13423-022-02136-7

Camos, V. (2008). Low working memory capacity impedes both efficiency and learning of number transcoding in children. *Journal of Experimental Child Psychology*, *99*(1), 37–57. https://doi.org/10.1016/j.jecp.2007.06.006

Campbell, J. (1995). Mechanisms of Simple Addition and Multiplication: A Modified Network-interference Theory and Simulation. *Mathematical Cognition*, *1*, 121–164.

Campbell, J. I. D., & Epp, L. J. (2004). An Encoding-Complex Approach to Numerical Cognition in Chinese-English Bilinguals. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *58*(4), 229–244. https://doi.org/10.1037/h0087447

Cerda, V. R., Grenier, A. E., & Wicha, N. Y. Y. (2019). Bilingual children access multiplication facts from semantic memory equivalently across languages: Evidence from the N400. *Brain and Language*, *198*, 104679. https://doi.org/10.1016/j.bandl.2019.104679

Chincotta, D., & Underwood, G. (1996). Mother tongue, language of schooling and bilingual digit span. *British Journal of Psychology*, *87*(2), 193–208. https://doi.org/10.1111/j.2044-8295.1996.tb02585.x

Chincotta, D., & Underwood, G. (1997). Speech Rate Estimates, Language of Schooling and Bilingual Digit Span. *European Journal of Cognitive Psychology*, *9*(3), 325–348. https://doi.org/10.1080/713752562

Chrisomalis, S. (2010). *Numerical Notation: A Comparative History*. Cambridge University Press.

Clayton, F. J., Copper, C., Steiner, A. F., Banfi, C., Finke, S., Landerl, K., & Göbel, S. M. (2020). Two-digit number writing and arithmetic in Year 1 children: Does number word inversion matter? *Cognitive Development*, *56*, 100967. https://doi.org/10.1016/j.cogdev.2020.100967

Cohen, J., Mac Whinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers*, *25*(2), 257–271.

Colomé, àngels, Laka, I., & Sebastián-Gallés, N. (2010). Language effects in addition: How you say it counts. *Quarterly Journal of Experimental Psychology*, *63*(5), 965–983. https://doi.org/10.1080/17470210903134377

Colomé, À. (2001). Lexical Activation in Bilinguals' Speech Production: Language-Specific or Language-Independent? *Journal of Memory and Language*, *45*(4), 721–736. https://doi.org/10.1006/jmla.2001.2793

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589–608. https://doi.org/10.1037/0033-295X.100.4.589

Comrie, B. (2013). Numeral bases (v2020.3). In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Zenodo. https://doi.org/10.5281/zenodo.7385533

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002

de Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. Psychology Press. https://doi.org/10.4324/9780203841228

De Visscher, A., & Noël, M.-P. (2014). The detrimental effect of interference in multiplication facts storing: Typical development and individual differences. *Journal of Experimental Psychology: General*, *143*(6), 2380–2400. https://doi.org/10.1037/xge0000029

De Vos, T. (1992). Tempo test rekenen (TTR)[Arithmetic number fact test]. *Nijmegen: Berkhout*.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*(1), 1–29. https://doi.org/10.1016/0010-0277(92)90030-l

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence. *Science*, *284*(5416), 970–974. https://doi.org/10.1126/science.284.5416.970

Del Maschio, N., & Abutalebi, J. (2019). Language Organization in the Bilingual and Multilingual Brain. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the Neuroscience of Multilingualism* (pp. 197–213). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119387725.ch9

Deloche, G., & Seron, X. (1982). From three to 3: A differential analysis of skills in transcoding quantities between patients with Broca’s and Wernicke’s aphasia. *Brain*, *105*(4), 719–733. https://doi.org/10.1093/brain/105.4.719

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, *82*(1), 64–81. https://doi.org/10.1348/2044-8279.002002

Dewaele, J.-M. (2007). Multilinguals’ language choice for mental calculation. *Intercultural Pragmatics*, *4*(3). https://doi.org/10.1515/IP.2007.017

Dijkstra, T. (2005). Bilingual Visual Word Recognition and Lexical Access. In *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). Oxford University Press. https://doi.org/10.1017/S0272263107210071

Dijkstra, T., & Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Wahl, A., Buytenhuijs, F., Halem, N. V., Al-Jibouri, Z., Korte, M. D., & Rekké, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Dotan, D., & Friedmann, N. (2018). A cognitive model for multidigit number reading: Inferences from individuals with selective impairments. *Cortex*, *101*, 249–281. https://doi.org/10.1016/j.cortex.2017.10.025

Dowker, A., & Nuerk, H.-C. (2016). Editorial: Linguistic Influences on Mathematics. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01035

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Duyck, W., & Brysbaert, M. (2002). What number translation studies can teach us about the lexico-semantic organisation in bilinguals. *Psychologica Belgica*, *42*(3), 151–175.

Duyck, W., & Brysbaert, M. (2004). Forward and Backward Number Translation Requires Conceptual Mediation in Both Balanced and Unbalanced Bilinguals. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 889–906. https://doi.org/10.1037/0096-1523.30.5.889

Duyck, W., Depestel, I., Fias, W., & Reynvoet, B. (2008). Cross-lingual numerical distance priming with second-language number words in native- to third-language number word translation. *Quarterly Journal of Experimental Psychology*, *61*(9), 1281–1290. https://doi.org/10.1080/17470210802000679

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of Acquisition Effects in Adult Lexical Processing Reflect Loss of Plasticity in Maturing Systems: Insights From Connectionist Networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1103–1123. https://doi.org/10.1037/0278-7393.26.5.1103

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology*, *71*(1), 43–51. https://doi.org/10.1111/j.2044-8295.1980.tb02728.x

Fiebach, C. J., Friederici, A. D., Müller, K., von Cramon, D. Y., & Hernandez, A. E. (2003). Distinct brain representations for early and late learned words. *NeuroImage*, *19*(4), 1627–1637. https://doi.org/10.1016/S1053-8119(03)00227-1

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: A unified JavaScript framework for online studies. *International Conference on Computational Social Science (Cologne)*. https://www.labvanced.com/

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*(3), 819–824. https://doi.org/10.1016/j.cognition.2008.04.007

Frenck-Mestre, C., & Vaid, J. (1993). Activation of number facts in bilinguals. *Memory & Cognition*, *21*(6), 809–818. https://doi.org/10.3758/BF03202748

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/10.1016/j.edurev.2013.05.003

Garcia, O., Faghihi, N., Raola, A. R., & Vaid, J. (2021). Factors influencing bilinguals' speed and accuracy of number judgments across languages: A meta-analytic review. *Journal of Memory and Language*, *118*, 104211. https://doi.org/10.1016/j.jml.2020.104211

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, *8*(3), 259. https://doi.org/10.1007/BF03174081

Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of Arithmetical Competencies in Chinese and American Children: Influence of Age, Language, and Schooling. *Child Development*, *67*(5), 2022–2044. https://doi.org/10.1111/j.1467-8624.1996.tb01841.x

Göbel, S. M., Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H.-C. (2014). Language affects symbolic arithmetic in children: The case of number word inversion. *Journal of Experimental Child Psychology*, *119*, 17–25. https://doi.org/10.1016/j.jecp.2013.10.001

Göbel, S. M., Watson, S. E., Lervåg, A., & Hulme, C. (2014). Children's Arithmetic Development: It Is Number Knowledge, Not the Approximate Number Sense, That Counts. *Psychological Science*, *25*(3), 789–798. https://doi.org/10.1177/0956797613516471

Grabner, R. H., Saalbach, H., & Eckstein, D. (2012a). Language-Switching Costs in Bilingual Mathematics Learning. *Mind, Brain, and Education*, *6*(3), 147–155. https://doi.org/10.1111/j.1751-228X.2012.01150.x

Grabner, R. H., Saalbach, H., & Eckstein, D. (2012b). Language-Switching Costs in Bilingual Mathematics Learning. *Mind, Brain, and Education*, *6*(3), 147–155. https://doi.org/10.1111/j.1751-228X.2012.01150.x

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

Greisen, M., Georges, C., Hornung, C., Sonnleitner, P., & Schiltz, C. (2021). Learning mathematics with shackles: How lower reading comprehension in the language of mathematics instruction accounts for lower mathematics achievement in speakers of different home languages. *Acta Psychologica*, *221*, 103456. https://doi.org/10.1016/j.actpsy.2021.103456

Grosjean, F. (2001). *The Bilingual's Language Modes*. Blackwell Publishing.

Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.

Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.

Hahn, C. G. K., Saalbach, H., & Grabner, R. H. (2017). Language-dependent knowledge acquisition: Investigating bilingual arithmetic learning. *Bilingualism: Language and Cognition*, *22*(1), 47–57. https://doi.org/10.1017/S1366728917000530

Hahn, C. G. K., Saalbach, H., & Grabner, R. H. (2019). Language-dependent knowledge acquisition: Investigating bilingual arithmetic learning. *Bilingualism: Language and Cognition*, *22*(1), 47–57. https://doi.org/10.1017/S1366728917000530

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). *Individual differences in non-verbal number acuity correlate with maths achievement*. *455*, 655–668. https://doi.org/10.1038/nature07246

Haspelmath, M., Dryer, M. S., Gil, D., & Comrie, B. (2005). *The World Atlas of Language Structures*. Oxford Univ. Press.

Hernandez, A. E. (2013). *The bilingual brain*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199828111.001.0001

Hirsh, K. W., Morrison, C. M., Gaset, S., & Carnicer, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition*, *6*(2), 117–128. https://doi.org/10.1017/S136672890300107X

Ifrah, G., & Bellos, D. (2000). *The universal history of numbers: From prehistory to the invention of the computer*. Wiley.

Imbo, I., Vanden Bulcke, C., De Brauwer, J., & Fias, W. (2014). Sixty-four or four-and-sixty? The influence of language and working memory on children's number transcoding. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00313

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277–288. https://doi.org/10.1016/j.actpsy.2007.06.003

Kempert, S., Saalbach, H., & Hardy, I. (2011). Cognitive benefits and costs of bilingualism in elementary school students: The case of mathematical word problems. *Journal of Educational Psychology*, *103*(3), 547–561. https://doi.org/10.1037/a0023619

Klaus, J., & Schriefers, H. (2019). Bilingual Word Production. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the Neuroscience of Multilingualism* (pp. 214–229). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119387725.ch10

Kochari, A. R. (2019). Conducting Web-Based Experiments for Numerical Cognition Research. *Journal of Cognition*, *2*(1), 39. https://doi.org/10.5334/joc.85

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed Numbers: Exploring the Modularity of Numerical Representations With Masked and Unmasked Semantic Priming. *Journal of Experimental Psychology: Human Perception and Performance*, 24. https://doi.org/10.1037/0096-1523.25.6.1882

Kovelman, I., Baker, S. A., & Petitto, L.-A. (2008). Age of first bilingual language exposure as a new window into bilingual reading development. *Bilingualism: Language and Cognition*, *11*(2), 203–223. https://doi.org/10.1017/S1366728908003386

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01795

Krinzinger, H., Gregoire, J., Desoete, A., Kaufmann, L., Nuerk, H.-C., & Willmes, K. (2011). Differential Language Effects on Numerical Skills in Second Grade. *Journal of Cross-Cultural Psychology*, *42*(4), 614–629. https://doi.org/10.1177/0022022111406252

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, *13*(3), 373–381. https://doi.org/10.1017/S136672891000009X

Lachelin, R., Marinova, M., Reynvoet, B., & Schiltz, C. (2023). Weaker semantic priming effects with number words in the second language of math learning. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001526

Lachelin, R., Rinsveld, A. van, Poncin, A., & Schiltz, C. (2022). Number transcoding in bilinguals—A transversal developmental study. *PLOS ONE*, *17*(8), e0273391. https://doi.org/10.1371/journal.pone.0273391

Lê, M.-L., & Noël, M.-P. (2021). Preschoolers' mastery of advanced counting: The best predictor of addition skills 2 years later. *Journal of Experimental Child Psychology*, *212*, 105252. https://doi.org/10.1016/j.jecp.2021.105252

Lê, M.-L. T., & Noël, M.-P. (2020). Transparent number-naming system gives only limited advantage for preschooler's numerical development: Comparisons of Vietnamese and French-speaking children. *PLOS ONE*, *15*(12), e0243472. https://doi.org/10.1371/journal.pone.0243472

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia*, *41*(14), 1942–1958. https://doi.org/10.1016/S0028-3932(03)00123-4

Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means* [Manual]. https://CRAN.R-project.org/package=emmeans

Lin, J.-F. L., Imada, T., & Kuhl, P. K. (2012). Mental Addition in Bilinguals: An fMRI Study of Task-Related and Performance-Related Activation. *Cerebral Cortex*, *22*(8), 1851–1861. https://doi.org/10.1093/cercor/bhr263

Lin, J.-F. L., Imada, T., & Kuhl, P. K. (2019). Neuroplasticity, bilingualism, and mental mathematics: A behavior-MEG study. *Brain and Cognition*, *134*, 122–134. https://doi.org/10.1016/j.bandc.2019.03.006

Lonnemann, J., & Yan, S. (2015). Does number word inversion affect arithmetic processes in adults? *Trends in Neuroscience and Education*, *4*(1), 1–5. https://doi.org/10.1016/j.tine.2015.01.002

Major, C. S., Paul, J. M., & Reeve, R. A. (2017). TEMA and Dot Enumeration Profiles Predict Mental Addition Problem Solving Speed Longitudinally. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02263

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115. https://doi.org/10.1017/S1366728903001068

Marinova, M., Georges, C., Guillaume, M., Reynvoet, B., Schiltz, C., & Van Rinsveld, A. (2021). Automatic integration of numerical formats examined with frequency-tagged EEG. *Scientific Reports*, *11*(1), 21405. https://doi.org/10.1038/s41598-021-00738-0

Marsh, L. G., & Maki, R. H. (1976). Efficiency of arithmetic operations in bilinguals as a function of language. *Memory & Cognition*, *4*(4), 459–464. https://doi.org/10.3758/BF03213203

Martinez-Lincoln, A., Cortinas, C., & Wicha, N. Y. Y. (2015). Arithmetic memory networks established in childhood are changed by experience in adulthood. *Neuroscience Letters*, *0*, 325–330. https://doi.org/10.1016/j.neulet.2014.11.010

Martini, S. (2021). *The influence of language on mathematics in a multilingual educational setting*. University of Luxembourg.

McClain, L., & Huang, J. Y. S. (1982). Speed of simple arithmetic in bilinguals. *Memory & Cognition*, *10*(6), 591–596. https://doi.org/10.3758/BF03202441

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, *44*(1–2), 107–157. https://doi.org/10.1016/0010-0277(92)90052-J

McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition*, *4*(2), 171–196. https://doi.org/10.1016/0278-2626(85)90069-7

McClung, N. A., & Arya, D. J. (2018). Individual Differences in Fourth-Grade Math Achievement in Chinese and English. *Frontiers in Education*, *3*, 29. https://doi.org/10.3389/feduc.2018.00029

Meeuwissen, M., Roelofs, A., & Levelt, W. J. M. (2003). Planning levels in naming and reading complex numerals. *Memory & Cognition*, *31*(8), 1238–1248. https://doi.org/10.3758/BF03195807

Miller, K. F., Smith, C. M., Zhu, J., & Zhang, H. (1995). Preschool Origins of Cross-National Differences in Mathematical Competence: The Role of Number-Naming Systems. *Psychological Science*, *6*(1), 56–60.

Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development*, *2*(3), 279–305. https://doi.org/10.1016/S0885-2014(87)90091-8

Ministère de l'Éducation Nationale. (2022, April 20). Languages in Luxembourg Schools. https://men.public.lu/en/themes-transversaux/langues-ecole-luxembourgeoise.html

Miura, I. T., Kim, C. C., Chang, C.-M., & Okamoto, Y. (1988). Effects of Language Characteristics on Children's Cognitive Representation of Number: Cross-National Comparisons. *Child Development*, *59*(6), 1445. https://doi.org/10.2307/1130659

Moeller, K., Shaki, S., Göbel, S. M., & Nuerk, H.-C. (2015). Language influences number processing – A quadrilingual study. *Cognition*, *136*, 150–155. https://doi.org/10.1016/j.cognition.2014.11.003

Moeller, K., Zuber, J., Olsen, N., Nuerk, H.-C., & Willmes, K. (2015). Intransparent German number words complicate transcoding – a translingual comparison with Japanese. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00740

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), Article 5109. https://doi.org/10.1038/2151519a0

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious Masked Priming Depends on Temporal Attention. *Psychological Science*, *13*(5), 416–424. https://doi.org/10.1111/1467-9280.00474

Naccache, L., & Dehaene, S. (2001). The Priming Method: Imaging Unconscious Repetition Priming Reveals an Abstract Representation of Number in the Parietal Lobes. *Cerebral Cortex*, *11*(10), 966–974. https://doi.org/10.1093/cercor/11.10.966

Negen, J., & Sarnecka, B. W. (2009). *Young children's number-word knowledge predicts their performance on a nonlinguistic number task*. https://escholarship.org/uc/item/1q03q75z

Negen, J., & Sarnecka, B. W. (2012). Number-Concept Acquisition and General Vocabulary Development. *Child Development*, *83*(6), 2019–2027. https://doi.org/10.1111/j.1467-8624.2012.01815.x

Notebaert, K., Pesenti, M., & Reynvoet, B. (2010). The neural origin of the priming distance effect: Distance-dependent recovery of parietal activation using symbolic magnitudes. *Human Brain Mapping*, *31*(5), 669–677. https://doi.org/10.1002/hbm.20896

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition*, *111*(2), 275–279. https://doi.org/10.1016/j.cognition.2009.02.002

Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition*, *82*(1), B25–B33. https://doi.org/10.1016/S0010-0277(01)00142-1

Nuerk, H.-C., Weger, U., & Willmes, K. (2004). On the Perceptual Generality of the Unit-Decade Compatibility Effect. *Experimental Psychology*, *51*(1), 72–79. https://doi.org/10.1027/1618-3169.51.1.72

Nuerk, H.-C., Weger, U., & Willmes, K. (2005). Language effects in magnitude comparison: Small, but not irrelevant. *Brain and Language*, *92*(3), 262–277. https://doi.org/10.1016/j.bandl.2004.06.107

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, *15*(2), 202–206. https://doi.org/10.1016/j.conb.2005.03.007

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503. https://doi.org/10.1126/science.1102085

Pitt, B., Gibson, E., & Piantadosi, S. T. (2022). *Exact number concepts are limited to the verbal count range*. *33*(3), 371–381. https://doi.org/10.1177/09567976211034502

Pixner, S., Zuber, J., Heřmanová, V., Kaufmann, L., Nuerk, H.-C., & Moeller, K. (2011). One language, two number-word systems and many problems: Numerical cognition in the Czech language. *Research in Developmental Disabilities*, *32*(6), 2683–2689. https://doi.org/10.1016/j.ridd.2011.06.004

Poncin, A., Rinsveld, A. V., & Schiltz, C. (2020). *Units first or tens first: How bilingualism affects two-digit number transcoding?* PsyArXiv. https://doi.org/10.31234/osf.io/sg7ea

Poncin, A., Van Rinsveld, A., & Schiltz, C. (2019). Units-first or tens-first: Does language matter when processing visually presented two-digit numbers? *Quarterly Journal of Experimental Psychology*, *73*(5), 726–738. https://doi.org/10.1177/1747021819892165

Power, R. J. D., & Dal Martello, M. F. (1990). The dictation of Italian numerals. *Language and Cognitive Processes*, *5*(3), 237–254. https://doi.org/10.1080/01690969008402106

Prior, A., Katz, M., Mahajna, I., & Rubinsten, O. (2015). Number word structure in first and second language influences arithmetic skills. *Frontiers in Psychology*, *6*(MAR), 266. https://doi.org/10.3389/fpsyg.2015.00266

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859–862. https://doi.org/10.3758/BF03192979

R Core Team. (2013). *R: A language and environment for statistical computing*. http://www.R-project.org/

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, *72*(2), 191–201. https://doi.org/10.1016/S0010-0277(99)00048-7

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic priming in number naming. *The Quarterly Journal of Experimental Psychology Section A*, *55*(4), 1127–1139. https://doi.org/10.1080/02724980244000116

Reynvoet, B., De Smedt, B., & Van den Bussche, E. (2009). Children's representation of symbolic magnitude: The development of the priming distance effect. *Journal of Experimental Child Psychology*, *103*(4), 480–489. https://doi.org/10.1016/j.jecp.2009.01.007

Rodic, M., Zhou, X., Tikhomirova, T., Wei, W., Malykh, S., Ismatulina, V., Sabirova, E., Davidova, Y., Tosto, M. G., Lemelin, J.-P., & Kovas, Y. (2015). Cross-cultural investigation into cognitive underpinnings of individual differences in early arithmetic. *Developmental Science*, *18*(1), 165–174. https://doi.org/10.1111/desc.12204

RStudio Team. (2020). *RStudio: Integrated development environment for r* [Manual]. http://www.rstudio.com/

Saalbach, H., Eckstein, D., Andri, N., Hobi, R., & Grabner, R. H. (2013). When language of instruction and language of application differ: Cognitive costs of bilingual mathematics learning. *Learning and Instruction*, *26*, 36–44. https://doi.org/10.1016/j.learninstruc.2013.01.002

Salillas, E., Barraza, P., & Carreiras, M. (2015). Oscillatory Brain Activity Reveals Linguistic Prints in the Quantity Code. *PLOS ONE*, *10*(4), e0121434. https://doi.org/10.1371/journal.pone.0121434

Salillas, E., & Carreiras, M. (2014). Core number representations are shaped by language. *Cortex*, *52*, 1–11. https://doi.org/10.1016/j.cortex.2013.12.009

Salillas, E., & Martínez, A. (2018). Linguistic Traces in Core Numerical Knowledge: An Approach From Bilingualism. In *Language and Culture in Mathematical Cognition* (pp. 173–196). Elsevier. https://linkinghub.elsevier.com/retrieve/pii/B9780128125748000080

Salillas, E., & Wicha, N. Y. Y. (2012). Early Learning Shapes the Memory Networks for Arithmetic: Evidence From Brain Potentials in Bilinguals. *Psychological Science*, *23*(7), 745–755. https://doi.org/10.1177/0956797612446347

Sasanguie, D., Defever, E., Van den Bussche, E., & Reynvoet, B. (2011). The reliability of and the relation between non-symbolic numerical distance effects in comparison, same-different judgments and priming. *Acta Psychologica*, *136*(1), 73–80. https://doi.org/10.1016/j.actpsy.2010.10.004

Sasanguie, D., & Reynvoet, B. (2014). Adults' Arithmetic Builds on Fast and Automatic Processing of Arabic Digits: Evidence from an Audiovisual Matching Paradigm. *PLOS ONE*, *9*(2), e87739. https://doi.org/10.1371/journal.pone.0087739

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., & Smedt, B. D. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3), e12372. https://doi.org/10.1111/desc.12372

Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology*, *12*(3), 281–300. https://doi.org/10.1111/j.2044-835X.1994.tb00635.x

Singmann, H. (2021). *Mixed Model Reanalysis of RT data* [Computer software]. https://cran.r-project.org/web/packages/afex/vignettes/afex_mixed_example.html

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *afex: Analysis of factorial experiments* [Computer software]. https://CRAN.R-project.org/package=afex

Spaepen, E., Coppola, M., Flaherty, M., Spelke, E., & Goldin-Meadow, S. (2013). Generating a lexicon without a language model: Do words for number count? *Journal of Memory and Language*, *69*(4), 496–505. https://doi.org/10.1016/j.jml.2013.05.004

Spelke, E. S., & Tsivkin, S. (2001a). Initial knowledge and conceptual change: Space and number. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (pp. 70–98). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.005

Spelke, E. S., & Tsivkin, S. (2001b). Language and number: A bilingual training study. *Cognition*, *78*(1), 45–88. https://doi.org/10.1016/S0010-0277(00)00108-6

Steiner, A. F., Banfi, C., Finke, S., Kemény, F., Clayton, F. J., Göbel, S. M., & Landerl, K. (2021). Twenty-four or four-and-twenty: Language modulates cross-modal matching for multidigit numbers in children and adults. *Journal of Experimental Child Psychology*, *202*, 104970. https://doi.org/10.1016/j.jecp.2020.104970

Steiner, A. F., Finke, S., Clayton, F. J., Banfi, C., Kemény, F., Göbel, S. M., & Landerl, K. (2021). Language effects in early development of number writing and reading. *Journal of Numerical Cognition*, *7*(3), 368–387. https://doi.org/10.5964/jnc.6929

Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual Word Recognition in a Sentence Context. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00174

van der Ven, S. H. G., Klaiber, J. D., & van der Maas, H. L. J. (2017). Four and twenty blackbirds: How transcoding ability mediates the relationship between visuospatial working memory and math in a language with inversion. *Educational Psychology*, *37*(4), 487–505. https://doi.org/10.1080/01443410.2016.1150421

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

van Marle, K., Chu, F. W., Li, Y., & Geary, D. C. (2014). Acuity of the approximate number system and preschoolers' quantitative development. *Developmental Science*, *17*(4), 492–505. https://doi.org/10.1111/desc.12143

Van Rinsveld, A., Brunner, M., Landerl, K., Schiltz, C., & Ugen, S. (2015). The relation between language and arithmetic in bilinguals: Insights from different stages of language acquisition. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00265

Van Rinsveld, A., Dricot, L., Guillaume, M., Rossion, B., & Schiltz, C. (2017). Mental arithmetic in the bilingual brain: Language matters. *Neuropsychologia*, *101*, 17–29. https://doi.org/10.1016/j.neuropsychologia.2017.05.009

Van Rinsveld, A., & Schiltz, C. (2016). Sixty-twelve = Seventy-two? A cross-linguistic comparison of children's number transcoding. *British Journal of Developmental Psychology*, *34*(3), 461–468. https://doi.org/10.1111/bjdp.12151

Van Rinsveld, A., Schiltz, C., Brunner, M., Landerl, K., & Ugen, S. (2016). Solving arithmetic problems in first and second language: Does the language context matter? *Learning and Instruction*, *42*, 72–82. https://doi.org/10.1016/j.learninstruc.2016.01.003

Van Rinsveld, A., Schiltz, C., Landerl, K., Brunner, M., & Ugen, S. (2016). Speaking two languages with different number naming systems: What implications for magnitude judgments in bilinguals at different stages of language acquisition? *Cognitive Processing*, *17*(3), 225–241. https://doi.org/10.1007/s10339-016-0762-9

Vander Beken, H., & Brysbaert, M. (2018). Studying texts in a second language: The importance of test type. *Bilingualism: Language and Cognition*, *21*(5), 1062–1074. https://doi.org/10.1017/S1366728917000189

Venkatraman, V., Siong, S. C., Chee, M. W. L., & Ansari, D. (2006). Effect of Language Switching on Arithmetic: A Bilingual fMRI Study. *Journal of Cognitive Neuroscience*, *18*(1), 64–74. https://doi.org/10.1162/089892906775250030

Volmer, E., Grabner, R. H., & Saalbach, H. (2018). Language switching costs in bilingual mathematics learning: Transfer effects and individual differences. *Zeitschrift Für Erziehungswissenschaft*, *21*(1), 71–96. https://doi.org/10.1007/s11618-017-0795-6

Wang, Y., Lin, L., Kuhl, P., & Hirsch, J. (2007). Mathematical and Linguistic Processing Differs Between Native and Second Languages: An fMRI Study. *Brain Imaging and Behavior*, *1*(3–4), 68–82. https://doi.org/10.1007/s11682-007-9007-y

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, *8*(3), 231–256. https://doi.org/10.1162/jocn.1996.8.3.231

Wicha, N. Y., Dickson, D. S., & Martinez-Lincoln, A. (2018). Arithmetic in the Bilingual Brain. In *Language and Culture in Mathematical Cognition* (pp. 145–172). Elsevier. https://doi.org/10.1016/B978-0-12-812574-8.00007-9

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251. https://doi.org/10.1016/0010-0285(92)90008-P

Xenidou-Dervou, I., Gilmore, C., van der Schoot, M., & van Lieshout, E. C. D. M. (2015). The developmental onset of symbolic approximation: Beyond nonsymbolic representations, the language of numbers matters. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00487

Xenidou-Dervou, I., van Atteveldt, N., Surducan, I. M., Reynvoet, B., Rossi, S., & Gilmore, C. (2023). Multiple number-naming associations: How the inversion property affects adults' two-digit number processing. *Quarterly Journal of Experimental Psychology*, 17470218231181367. https://doi.org/10.1177/17470218231181367

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Ziegler, J. C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical Words are Read Differently in Different Languages. *Psychological Science*, *12*(5), 379–384. https://doi.org/10.1111/1467-9280.00370

Zuber, J., Pixner, S., Moeller, K., & Nuerk, H.-C. (2009). On the language specificity of basic number processing: Transcoding in a language with inversion and its relation to working memory capacity. *Journal of Experimental Child Psychology*, *102*(1), 60–77. https://doi.org/10.1016/j.jecp.2008.04.003

1 STUDY 3

Weaker semantic priming effects with number words in the second language of math learning

Rémy Lachelin[1], Mila Marinova[1], Bert Reynvoet[2] and Christine Schiltz[1]

[1]University of Luxembourg [2]KU Leuven

Published:

Lachelin, R., Marinova, M., Reynvoet, B., & Schiltz, C. (2023). Weaker semantic priming effects with number words in the second language of math learning. *Journal of Experimental Psychology: General. https://doi.org/10.1037/xge0001526*

1.1 Abstract

Bilinguals' exact number representations result from associations between language-independent Indo-Arabic digits (*i.e.* "5"), two verbal codes (*i.e.* "fünf" and "cinq") and a common, largely overlapping semantic representation. To compare the lexical and semantic access to number representations between two languages, we recruited a sample of balanced highly proficient German-French adult bilinguals. At school, those bilinguals learned mathematics in German for 6 years (LM1) and then switched to French (LM2) in 7th grade (12 years old) until 13th grade (for 7 years). After the brief presentation of primes (51ms) consisting of Indo-Arabic digits or number words in German or French, an Indo-Arabic digits target had to be read in either German or French in an online study. Stimuli were numbers from 1 to 9 and we varied the absolute distance between primes and targets from 0 (*i.e.* 1 - 1) to 3 (1 - 4) (as in Reynvoet et al., 2002). The priming distance effect (PDE) was used to measure the strength of numerical semantic association. We find comparable PDEs with Indo-Arabic digits and German number word primes, independently from the target naming language. However, we did not find a clear PDE with French number word primes, neither when naming targets in German, nor in French. The weaker PDE from LM2 compared to LM1 primes is interpreted as a weaker lexico-semantic association of LM2 number words. These results indicate a critical role of the first language of math learning and further emphasize the role of language in processing numbers. They might have important implications for designing bilingual school curricula.

Keywords: Numerical cognition, priming distance effect, bilingualism, language of mathematical learning

1.2 Public Significance Statement:

This study demonstrated a cognitive cost for highly proficient bilinguals when processing the meaning of numbers in a second language. The cost was observed even though bilinguals had attended math classes in the second language (French) during 7 school years, following math acquisition in the first language (German) over a period of 6 school years. Our findings indicate that sequential bilingual school curricula imply a cost for processing numbers in the second language. Given the hierarchical nature of math education and its fundamental importance for later academic and professional achievement, this cost should ideally be acknowledged and addressed to assure optimal learning outcomes.

2 Introduction

Human beings have non-symbolic and symbolic representations of numerosities. Non-symbolic number representations (*i.e.*●●●●●) are approximate and functional very early in the cognitive development (Barth et al., 2003; Halberda et al., 2008; F. Xu & Spelke, 2000). On the other hand, symbolic representations such as English number words (*i.e.* "five") and Indo-Arabic digits (*i.e.* "5") are precise and acquired later in development. The acquisition of number words promotes precise numerical representation as sustained by developmental studies (Negen & Sarnecka, 2009, 2012; Wynn, 1992b) and cross-linguistic studies on languages with restricted number words (Frank et al., 2008; Pica et al., 2004; Pitt et al., 2022; Spaepen et al., 2013). Number words' and Indo-Arabic digits' semantic representations are associated with common numerical features (*e.g.* magnitude, order or parity, Koechlin et al., 1999; Marinova et al., 2021). Moreover, the development of number semantic representations predicts later mathematic performances both when considering number words (Desoete et al., 2012; M.-L. Lê & Noël, 2021; Major et al., 2017; van Marle et al., 2014) and Indo-Arabic digits (Göbel, Watson, et al., 2014; M. Schneider et al., 2017). Yet, for bilinguals different sets of number words exist in their respective languages, in contrast to Indo-Arabic digits, which are in use across numerous languages and writing systems (Ifrah & Bellos, 2000). Therefore, bilinguals might show different strengths of association between number words of the different languages, Indo-Arabic digits and their semantic representations. The strength of association possibly depends on the language of mathematical education and might in turn influence mathematical performances (Van Rinsveld et al., 2017). The present study aims to investigate how the language of learning mathematics shapes lexical and semantic representations of number words of proficient bilinguals, which is of particular interest for bilingual school curricula.

2.1 Bilingual arithmetic and transcoding

Studies on bilinguals solving arithmetics or doing transcoding tasks (*i.e.* involving the conversion of a number between either its non-symbolic, verbal, or visual form) highlight the importance of the language in which mathematics are learned. Those studies reveal a cognitive cost, hence a worse performance on the same task when done in the less dominant language, or the language not used for formal math acquisition or ad-hoc arithmetic training. This cost can be measured as slower reaction times or more errors. In a seminal study by Spelke and Tsivkin

(2001), bilingual Russian dominant (L1) participants who later learned English (L2) were trained to solve arithmetic either in their L1 or L2. A consequent test conducted in the L1 and L2 resulted in a cost for solving arithmetics in the untrained language which was independent of the testing language (*i.e.* L1 or L2). Therefore, independently from language dominance (*i.e.* L1 or L2), participants performed worse when switching between learning and testing language: a language switching cost (LSC). LSCs have been measured behaviourally (Dehaene et al., 1999; Hahn et al., 2017; Saalbach et al., 2013; Volmer et al., 2018), as well as with fMRI and EEG neuroimaging methods indicating different brain activity when solving problems in an untrained language compared to the trained one (Grabner et al., 2012b; Venkatraman et al., 2006). Hence, LSC might have important consequences on how bilinguals learn language-dependent arithmetic facts. Bernardo (2001) investigated students with L1 dominant Philippino, who learned mathematics in English at school, indicating a cost for solving arithmetic in Philippino compared to English. These results suggest a critical role of the language of learning mathematics (LM) for arithmetic facts consolidation. Solving arithmetic in the L1 compared to a different LM also elicits distinct EEG responses (Salillas & Wicha, 2012, but see Cerda et al., 2019 and; Martinez-Lincoln et al., 2015). Hence, independent of the L1, those studies suggest a benefit for solving arithmetic in the LM.

The language-related cost arising during arithmetic might partially originate in the more elementary process of transcoding, which is thought to be a sub-process involved in solving arithmetic. For example, when solving "7 X 6 = ? ", the results could involve the passage from a visual to a verbal code, as suggested by correlations between reaction times to solve arithmetic and transcode numbers (Clayton et al., 2020; Steiner, Banfi, et al., 2021). Furthermore, both arithmetic and transcoding tasks reveal costs when performed in the less dominant or untrained language. For bilingual participants who followed the Luxembourgish school system where mathematics are thought first in German for six years (LM1) and then in French for seven years (LM2), slower response times and more errors for complex arithmetic are found for LM2 compared to LM1, even in adults (Van Rinsveld et al., 2015). The LM2 cost for French was further replicated in a second cross-sectional study for the more elementary task of transcoding two-digit Arabic numbers, also until adulthood (Lachelin et al., 2022; see also Garcia et al., 2021 for complementary results in a meta-analysis).

In sum, these studies reveal language-specific costs during arithmetic and transcoding tasks in bilinguals. However, it remains unknown from which specific processing level those costs arise when bilinguals are dealing with numbers. For example, in the case of German

(LM1) French (LM2) bilinguals, the cost in transcoding might be explained uniquely by lexical retrieval, *i.e.,* retrieving that "5" is "fünf" would be faster than "cinq". Or additional costs might be due to later weaker semantic associations. To address this open question, we used the priming distance effect as experimental paradigm (see section *Distance effect*) and relied on the triple code model as a theoretical framework (see section *Bilingual Triple Code Model*) to precisely locate the levels of language-specific costs during number processing in highly proficient bilinguals (see section *Heterogeneity in Bilingualism*).

2.2 Distance effect

The distance effect refers to a decrease in participant's performance when required to compare two numbers as the absolute difference between two numbers is reduced (*e.g.,* 5 vs 6 compared to 5 vs 9). It is commonly used to assess the semantic relation between numbers (Moyer & Landauer, 1967) and reveals activation of number semantics more generally. The distance effect can also be observed in priming paradigms: the priming distance effect (PDE). In this paradigm, the prime modulates reaction times as a function of the distance between prime and target, so that closer numerical distances (*i.e.* prime = 4, target = 5) elicit faster responses than distant pairs (*i.e.* prime = 2, target = 5) (Koechlin et al., 1999; Naccache et al., 2002). Developmentally, PDEs with Indo-Arabic digits as primes are already found in 1st graders, and remain stable for older age groups (Reynvoet et al., 2009). Remarkably the PDE can be elicited from primes presented as Indo-Arabic digits as well as number words (Reynvoet et al., 2002), thus allowing to test the semantic activation with number words in different languages. Despite the use of very fast and masked primes (*i.e.* 43 ms), these modulate both reaction times and cerebral responses as a hallmark of a distance effect (Koechlin et al., 1999; Naccache & Dehaene, 2001; K. Notebaert et al., 2010). PDEs are also observed when measuring voice onset times in experiments where the targets have to be named (Reynvoet & Brysbaert, 1999), thus allowing to compare responses in different languages with the same paradigm.

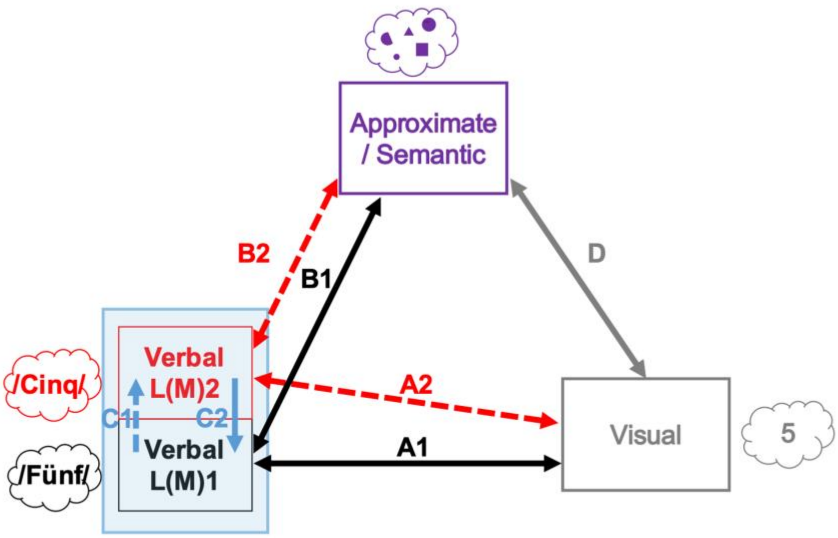
To test number semantic associations in multilinguals' different languages, number word translation PDE paradigms have been used. Thus Duyck et al. (2008) investigated Dutch (L1) - English (L2) - French (L3) speakers with L2 number word primes and L1 or L3 number word targets. The task was either to read the targets in the language they were written in (within language) or to translate them. When the same numerosity was presented as prime and target (*i.e.* in repetition priming trials, 2 in the examples) mean voice onset times were faster than with non-repeated primes with both forward (L1 targets translated to L3, *i.e.* prime = "two", target = "twee", response = /deux/) and backward translation (L3 to L1, *i.e.* prime = "two", target = "deux", response = /twee/). Moreover, PDE was observed when naming in L1 (both with L1 and L3 target number words, hence after a backward translation, *i.e.* prime = "two", target = "vijf" or "cinq ", response = /vijf/ or /vijf/). In contrast, no PDE was observed when naming in L3 (both with L3 and L1 targets, hence after a forward translation, *i.e.* prime = "two", target = "viijf " or "cinq ", response = /cinq/ or /cinq/). The interpretation was that backward translations have a stronger lexico-semantic association than forward translations. However, this study did not allow to compare lexico-semantic associations within each language, since the language of prime and target number words systematically differed. Duyck & Brysbaert (2002) partially addressed this question by presenting Dutch (L1) - French (L2) bilinguals with Indo-Arabic digits primes and L1 or L2 number word targets which had to be named in the presented language or translated. Again, PDEs were observed when naming in L1 (both with L1 and backward translated L2 target number words, *i.e.* prime = 2, target = "vijf" or "cinq" , response = /vijf/ or /vijf/) but they were absent in the group instructed to name in L2 (both with L2 and forward translated L1 target number words, *i.e.* prime = 2, target = "cinq" or "vijf", response = /cinq/ or /cinq/). Repetition priming, *i.e.* when the prime is the same number as the target, was stronger when targets had to be translated (*i.e.* prime = 2, target = "twee" or "deux ", response = /deux/ or /twee/) than named (*i.e.* prime = 2, target = "twee" or "deux ", response = /twee/ or /deux/).

These studies demonstrate that the PDE paradigm can be used to assess the lexico-semantic associations of numbers in bilinguals. However, they did not probe whether number words in L1 and L2 automatically elicit semantic activations when presented as prime briefly before Indo-Arabic digits have to be named. Furthermore, the participants in the above-mentioned studies were not balanced bilinguals since they acquired the L2 lately (*i.e.* as 10 year-olds) and L2 was not a language of learning mathematics for them.

2.3 Bilingual Triple Code Model

The triple code model (TCM) (Dehaene, 1992) synthesises the neurocognitive modular organisation between number words (verbal), Indo-Arabic digits (visual), and abstract semantics. The TCM's verbal code is part of general-purpose language abilities. The transcoding routes between verbal and visual are asemantic according to the TCM, meaning that the access to number's abstract code is not required. However, PDE experiments described previously suggest an automatic semantic co-activation from Indo-Arabic and number word primes. Those experiments also indicate that PDE activation of the semantic representation depends on prime notation (Koechlin et al., 1999). In the original formulation of the TCM (Dehaene, 1992) number semantic is activated only by the most dominant language of bilinguals. However, another possibility is the existence of language-specific parallel, but distinct, semantic associations. In this case, the gradient of semantic activation spread might vary not only with notation, but also with the languages of a bilingual.

Figure 1 *Bilingual Triple Code Model (BTCM) representing lexico-semantic and lexico-visual associations between each language's code represented with black and red arrows.*



Notes: L(M)1 = first Language (of Mathematical) learning, L(M)2 = second Language (of Mathematical) learning. Blue unidirectional arrows indicate translations between the two languages existing for the verbal code: forward translation from L(M)1 to L(M)2 (C1); backward translation from L(M)2 to L(M)1 (C2). Dashed arrows indicate weaker associations compared to full arrows. The arrows correspond to: bidirectional lexico-visual associations with L(M)1 (A1), bidirectional lexico-visual associations with L(M)2 (A2), semantic access from - and to - L(M)1 (B1), semantic access from - and to - L(M)2 (B2), independent semantic access from - and to - the visual code (D).

Here, we propose a rewriting of the TCM onto a bilingual triple code model (BTCM), such that each language-specific verbal code would have parallel bidirectional lexico-visual (see A1 and A2 in Figure 1) and lexico-semantic (B1 and B2) associations starting from each language-specific verbal codes; see Figure 1. Within the verbal code, a direct lexico-lexical connection between the language-specific verbal codes would also be available for number word translation, (C1 and C2 in Figure 1). From the literature we know that translation is easier from L2 to L1 (backward) than from L1 to L2 (forward), which is also predicted in the RHM (Kroll et al., 2010). From this BTCM framework, we can therefore compare each verbal code's specific lexical association, lexico-visual and lexico-semantic association, or activation. Hence, the strength of association of the verbal codes might differ between each of the bilinguals' languages.

The strength of association (dashed compared to full lines in Figure 1) could in part be determined by general language factors such as balanced bilingualism and L2 proficiency (*i.e.* Garcia et al., 2021). More specific factors affecting the strength of association are the language of (math) training (as for LSC, *i.e.,* Saalbach et al., 2013) or the language of learning mathematics (LM, Van Rinsveld et al., 2015). Depending on their specific configurations, these factors and their interactions could lead to weaker associations between verbal, visual and/or semantic processing levels and entail corresponding costs. Hence weaker L(M)2 associations are expected for unbalanced bilinguals as well as for bilinguals with low proficiency and/or less math training in L(M)2.

Alternative models accounting for transcoding in bilinguals include a version of the encoding complex model (ECM, Campbell & Epp, 2004) and a bilingual encoding complex model (Bernardo, 2001). There are three main differences between the proposed BTCM and ECM. First, while for the ECM the strength of associations between formats and languages depends on tasks and training (encoding-retrieval integration), the BTCM introduces age or order of acquisition as a factor such that earlier acquired languages have stronger associations. This point is relevant in practice with regards to bilingual education. Second, the BTCM assumes that both languages are integrated into a single lexicon rather than two separated ones. Third, the ECM does not include translations from one language to another and asymmetries regarding the strength of associations (*i.e.,* Figure 1, C1 and C2). An interesting connectionist model has also been proposed by (Duyck & Brysbaert, 2004). In this model, each lexicon of the different languages has different degrees of overlap of connections with its corresponding semantic representation, similarly as in the language general BIA+ model (Dijkstra & van Heuven, 2002). Note that in these models, differently to the original TCM proposition, the number semantics rather than being a separate system might emerge from the associations between numbers. This is also proposed by the "discrete semantic system", suggesting that the distance effect results from the semantic network between the numbers rather than from a separate semantic system as in the TCM (Krajcsi et al., 2016). Since so many interacting factors might contribute to a language cost, it is particularly important to understand the mechanisms and relevance of language proficiency in bilinguals.

2.4 Heterogeneity in bilingualism

Bilinguals[14] represent more than half of the world's population (Grosjean, 2008). They are proficient in two languages, an L1 and an L2, whose configurations can be very heterogeneous across subjects. For example, L1 and L2 proficiency can range between balanced and unbalanced bilinguals, while L2 proficiency can range between high and low (de Groot, 2011). The L2 proficiency depends on different factors such as the age of acquisition, language exposure, or L1 and L2 linguistic similarities, which in turn influences the organisation of the brain (Del Maschio & Abutalebi, 2019; Hernandez, 2013; Klaus & Schriefers, 2019). L1 and L2 are activated in parallel during comprehension and production (À. Colomé, 2001; Dijkstra, 2005; Marian & Spivey, 2003). This concurrent activation is controlled by top-down pre-frontal inhibitory mechanisms (Abutalebi, 2008; Green, 1998). The strength of top-down inhibition mechanics depends on both L2 and L1 proficiency, such that balanced bilinguals should have comparable inhibition strengths for both L1 and L2 (Costa & Santesteban, 2004), while bilinguals with high L2 proficiency have a stronger inhibition than those with a low L2 proficiency (de Groot, 2011). In addition, differences between L1 and L2 strength of activation might also occur at different language processing stages such as lemma, lexical or semantic (Kroll et al., 2010). On a theoretical level, weaker L2 compared to L1 activations affecting those different stages are predicted by several psycholinguistic models of bilingual language production and comprehension.

In sum language proficiency is an important marker of how languages are stored in the bilingual's brain. The ideal sample to study (numerical) cognition in bilinguals is thus composed of balanced, highly proficient bilinguals that have formally acquired both languages in school and that have grown up in an environment systematically exposing the individuals to both languages in a similar manner.

[14] Herein we will use and describe the specific case of bilingualism which is a subgroup for the more general term multilingualism (proficiency in multiple languages).

2.5 Present study

With the present study, we aim to measure lexical and semantic access from number words in a balanced bilingual sample that sequentially acquired mathematics in a first language (LM1) and then in a second language (LM2). We sampled adults who followed the Luxembourgish public schools where mathematics are learned in German (henceforth LM1) from 1st until 6th grade (about 12 years old). From 7th grade until the end of obligatory school (19 years old) the language to learn mathematics switches to French (henceforth LM2), thus resulting in highly proficient German-French bilinguals (*Ministère de l'Éducation Nationale*, 2022). To measure both lexical retrieval and semantic access from number words we implemented a PDE paradigm as in Reynvoet et al. (2002), using German and French number words as primes and Indo-Arabic digits as targets. Number words have a high degree of semantic overlap across languages (*i.e.* magnitude, order or parity). Visual Indo-Arabic digits constitute an additional association with both L1 and L2 number-word lexicons and semantics. This allowed us to measure the strength of priming through number words in both languages on Indo-Arabic digits, which can also be named in both languages. We defined the following hypotheses: first, we expected an LM2 cost for lexical retrieval, with slower voice onset times for Indo-Arabic digit naming in the LM2 than in the LM1, as suggested by previous studies (Garcia et al., 2021; Lachelin et al., 2022). Second, we expected weaker LM2 lexico-semantic associations, which would be reflected by weaker PDEs with LM2 than LM1 number word primes (Dijkstra & van Heuven, 2002; Kroll et al., 2010).

The study was implemented on Labvanced, a web-based platform (Finger et al., 2017). Previous replications of studies on numerical cognition have shown that even masked priming studies can be implemented on web-based platforms (Kochari, 2019).

3 Methods

3.1 Participants

A total of 39 participants completed the experiment in an exchange for 5 euros voucher. Seven participants were excluded because French was reported as the most proficient language. None of the participants reported antecedents of dyscalculia, dyslexia, or epilepsy. Hence, the final sample was composed of 32 participants (M_{age} = 23.6 years, *SD = 6.1* years, gender reported: 26 females, 6 males, 0 other). The sample reflected Luxembourg's multilingualism, with the participants reporting knowing on average M = 4.8*(0.8)* languages and all participants speaking Luxembourgish, German, French, and English.

The sample's average age of acquisition and frequency of use are described in Table 1 for the five most frequently reported languages. Note that linguistically speaking, Luxembourgish can be considered a German dialect (Martini, 2021), with number words being orthographically, phonologically, and morphologically very similar to German.

Table 1

Age of Acquisition (AoA indicated in years) and Frequency of Use of the language (Frequency of use)

	Luxembourgish	German	French	English	Portuguese
AoA	2.13 *(2.27); 13**	4.9 *(2.10); 2**	7.06 *(1.72); 0**	12.72 *(1.46); 0**	0.7 (1.06); 6*
Frequency of use	4.87 *(0.5)*	4.09 *(0.69)*	3.87 *(0.75)*	3.74 *(0.85)*	4.9 (0.32)
N	31	32	32	32	10

Note: AoA is reported with responses 0 included); * number of participants with response = 0. Frequency of use: 5 = daily, 4 = weekly, 3 = monthly, 2 = yearly and 1 = never (1 was not answered for these languages). N = number of participants reporting those languages. Standard deviations in parenthesis.

The specific AoA of language in which mathematics were learned was reported earlier for German (6.4*(1.7)* years old) than French (12.0*(2.3)* years old). This is fully in line with the Luxembourgish school curriculum where all topics are taught in German from 1st (6 years old) to 10th grade, except for mathematics, which is taught in French from 7th grade onwards (12 years old). 25 participants reported using their most proficient language (German/Luxembourgish for the majority) to solve different types of arithmetic problems. From 11th grade onwards, all topics are taught in French. This results in highly proficient German-French bilinguals (*Languages in Luxembourg Schools*, 2021). Therefore, in the following analyses, the Language of Learning Mathematics (LM) will be considered as a factor, with German being the first language of learning mathematics (LM1) and French the second (LM2).

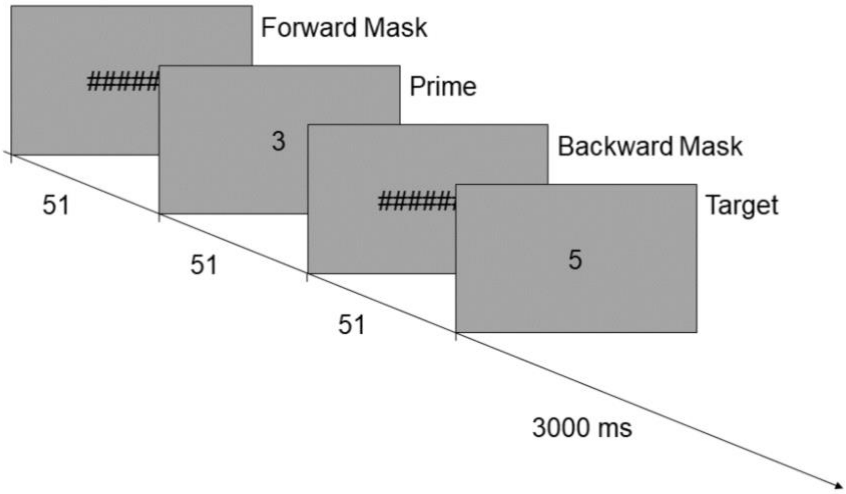
The Ethical Review Panel approved the experimental protocol at the University of Luxembourg (ERP 21-005 OnBiNNPri). Before undertaking the experiment, participants gave informed consent.

3.2 Priming Distance Effect (PDE) Task

Participants were presented with a masked priming task similar to the one used by Reynvoet et al. (2002). Both masks and primes lasted 51 ms. The masks were controlled to visually overlap the longest prime (*i.e.* "SIEBEN"). After the backward mask, the target was presented for 2500 ms, at which the microphone recording started, see Figure 2. The participant's task was to name the target, which was an Arabic digit for all trials. Prime awareness was asked at the end of the study but due to a technical error this response was not recorded.

The masks and stimuli were in black and were programmed to appear in the centre of a grey screen. All stimuli were presented within a 6 X 2 visual angles text box in the middle of the screen. Visual angle self-calibration was possible thanks to Labvanced's built-in feature requiring the participants to adjust the distance from the screen and calibrate the screen size with a standard-sized credit card at the beginning of the experiment. The participant saw an adjustable rectangle on the screen that could be adjusted with the mouse to match the size of the card.

Figure 2: *Time-line of a trial with an Indo-Arabic digits as a prime*



Note. Other trials included German or French number words as primes. The participant's task was to name the target, which was always an Indo-Arabic digit, in either German or French (blocked).

Participants' verbal responses to the targets, measured by the Voice Onset Times (VOT) serve as the dependent measure. The VOT were encoded using CheckVocal (Protopapas, 2007) by automatic voice onset detection. Then, an external experimenter, naïve to both the hypothesis and primes, visually and auditorily checked each recording. Manual adjustments were made whenever necessary. For instance, to correct the VOT for number words starting with fricatives (*i.e.* /vier/ or /deux/), and to identify any additional noise (e.g., mispronunciations, recording errors, etc..).

3.3 Stimuli

All the stimuli (targets and prime) were numbers ranging between 1 and 9, depicted as Indo-Arabic digits or number words. Primes varied in notations: Indo-Arabic digits (*i.e.* 5), German number words (*i.e.* FÜNF), and French number words (*i.e.* CINQ). The targets were always Indo-Arabic digits. Thus, both languages were retrieved from the same Indo-Arabic digit depending on the experimental condition. The *distances* between Prime Target pairs (*i.e.* $\mathrm{absolute}(\mathrm{target} - \mathrm{prime})$) were restricted to 0, 1, 2, and 3. That is, the distance 0 represents repetition priming since the same number value is presented as prime and target.

To avoid statistical prediction strategies each Indo-Arabic digit from 1 to 9 was equally frequent within each condition's target. To achieve this we had to balance the prime-target pairs, for example for the items corresponding to the distance 1, the prime target pairs 2-1 and 8-9 were presented twice, see S1 Table 1. Additionally, 18 trials with a "filler" prime (*i.e.* ######) were added to have a baseline of non-primed number naming. In sum, each of the two conditions contained 234 experimental trials (72 pairs with prime = number word German, 72 prime = number word French, 72 prime = digit, 18 filler trials). The different notation prime and target pairs were randomly presented within each condition, while naming languages were blocked.

3.4 Procedure

Participants were recruited via mailing lists and social media targeting university students with at least ten years of schooling in Luxembourg by sharing a link to the experiment hosted on the web-based platform Labvanced (Finger et al., 2017). Hence the experiment ran on the participant's personal computers at home. The participants were required to be in a quiet room where they would not be disturbed or distracted for the duration of the experiment. Each participant was randomly assigned to a German or French language starting condition (15 participants started in German, 17 in French). Before the task, the participant answered a short 13 item questionnaire about demographics, self-reported language use, and language for number processing (described above). The questionnaire was followed by written instructions and translated according to the starting language condition. This manipulation was done to balance the language mode before starting the experiment across the sample (see Grosjean, 2001). Then the task started with the condition where all targets had to be named in the starting language. Between each language block, participants could take a short break. The same stimuli set (prime-target pairs) was presented for both blocks, but their order of presentation was randomized before each block. The experiment lasted about 30 to 35 minutes. At the end of the experiment, each participant could indicate their contact information to receive their compensation.

3.5 Data analyses

Data were analysed using linear mixed models. All analyses were done with R (RStudio Team, 2020) and the following packages: for the linear mixed models afex (Singmann et al., 2020), which relies on lme4 (Bates et al., 2015). Follow-up analyses were computed with emmeans (Lenth, 2021), and graphs were drawn with ggplot2 (Wickham, 2016). Voice onset times (VOT) were used as a dependent variable.

3.6 Transparency and Openness

Data and R scripts to reproduce the results are available at: *https://osf.io/j8vzu/*.

4 Results

4.1 Task descriptive

From the 32 participants, a total of 14972 voice onset times, (VOT) were measured (a few participants who quit the experiment at maximum 32 trials before the end of the experiment due to program error were included). VOTs were filtered according to the three following criteria: first trials marked as mispronunciations, failed or unintelligible recordings (0.59% of the total). From these (i.e., 88/14972 trials), only 14 were mispronunciations and 3 were responses in English, hence too few for conducting further meaningful analyses on accuracy. Second, by applying highpass (300 ms) and lowpass (1500 ms) filters on the VOT (0.27%). Third, we excluded each VOT exceeding 3 standard deviations from individual means to remove outliers (1.50%). In sum, 2.36 % of the initial total trials were filtered out (ending with 14619 trials to analyse).

4.2 Filler prime

We analysed the trials with the "filler" primes (*i.e.* ######), that is the trials corresponding to digit number naming without any number priming. For these trials, we observed faster VOT in the block where the target had to be named in German (642*(89) ms*) than in French (665*(101) ms*) (paired t-test: $t(31) = -3.07$, $p = .004$). This result indicates a cost for lexical access from Indo-Arabic digits to the corresponding number word in the LM2 (French) compared to the LM1 (German).

4.3 Linear Mixed Model

Linear mixed models (LMM) were applied with: *Distance* (0, 1, 2, 3), *Prime Notation* (digit, German number word, and French number word), and target *Naming Language* (German or French) as fixed factors. Random slopes and intercepts were modelled to adjust for differences between the different *Target*'s number word length in both languages. Random intercepts for each *Subject* were also included in the model to take into account individual differences in VOT. The following maximal model (A) was defined a priori (Barr et al., 2013). Because the model (A) did fit and did not present any problems such as singularity we did not need to select or remove terms. All degrees of freedom of the following analyses were obtained by the Satterthwaite approximation method, comparing the full model against the model without the effect (Singmann et al., 2020). The R syntax of the main model was as in (A):

$$(A)\ VOT \sim Distance * PrimeNotation * NamingLanguage + (NamingLanguage|Target) + (NamingLanguage|Subject)$$

The main LMM resulted in a main effect of distance ($F(3,13382.14) = 140.07, p < .001$), prime notation ($F(2,13382.14) = 70.94, p < .001$), and target naming language ($F(1,17.06) = 6.10, p < .05$). Prime notation interacted with distance ($F(6,13382.14) = 9.83, p < .001$) and with target naming language ($F(2,13382.14) = 3.06, p = .05$). Since the main LMM showed a main effect of prime notation, we decided to conduct two separate LMMs for digits and number words. Separate analyses are also justified theoretically since both notation formats have different underlying cognitive processes (see Reynvoet et al., 2002). The same random effects by targets and participants were maintained as for the main LMM as described above. Table 2 depicts the VOT for each distance by prime notation and target naming language.

Table 2 *VOT for each distance and different primes*

Prime's notation:		Digit		NW German		NW French			#######	
		Target's naming language:								
		German	French	German	French	German	French		German	French
Distance	0	607*(106)*	628*(116)*	640*(110)*	670*(120)*	652*(101)*	657*(118)*	No prime condition:		
	1	650*(106)*	675*(107)*	662*(103)*	687*(104)*	668*(96)*	689*(116)*		642*(60)*	666*(65)*
	2	653*(95)*	681*(109)*	670*(104)*	694*(107)*	664*(99)*	687*(108)*			
	3	659*(95)*	687*(109)*	671*(103)*	698*(105)*	667*(99)*	695*(111)*			
Mean:		642*(103)*	668*(113)*	661*(105)*	687*(110)*	663*(99)*	682*(114)*		642*(60)*	666*(65)*

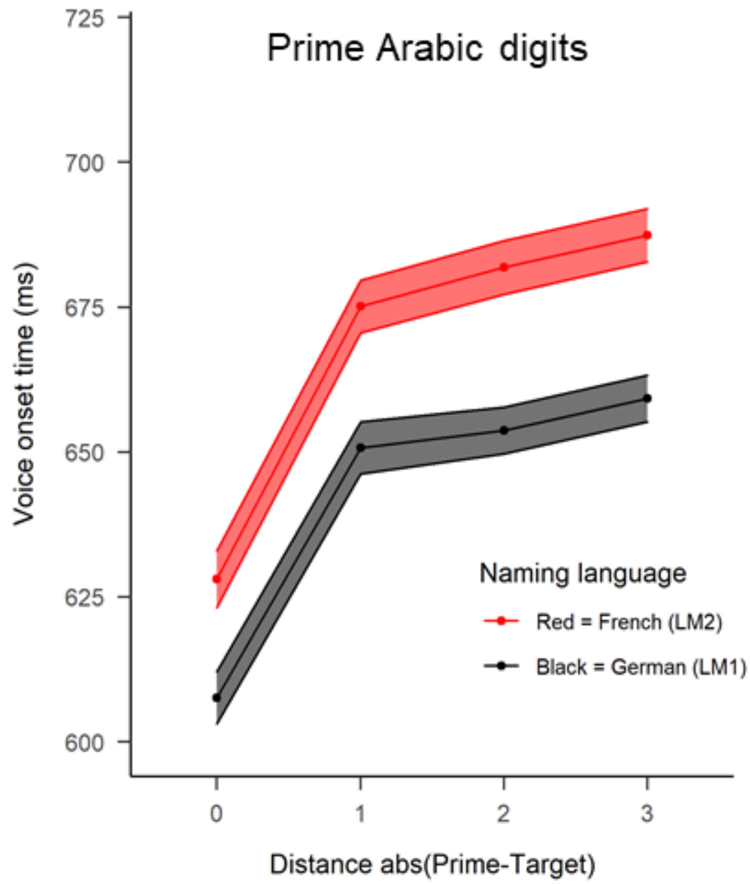
Notes. Average VOT in milliseconds (SD in parenthesis) from data aggregated by distance, prime notation, and target number naming language. NW = Number Words.

4.4 LMM by prime notations

4.4.1 Indo-Arabic digits

The LMM applied on the condition with Indo-Arabic digits as primes was as in formula (A), but without the main effect of *Prime Notation* format. This LMM yielded a main effect of the target's naming language ($F(1{,}19.90) = 6.76, p < .05$), with slower VOT when naming the targets in French than in German. This indicates a lexical retrieval cost for the LM2 compared to the LM1, as it was already visible in the *filler prime* condition. Furthermore, we also found a significant main effect of distance ($F(3{,}4420.31) = 106.69, p < .001$). Posthoc pairwise decomposition indicated first a significant repetition priming effect, as reflected by faster VOT for the distance 0 (*i.e.* same prime and target) than distance 1 ($t(4420.23) = -12.84, p_{holm} < .001, estimate = -44.81, SE = 3.49$), Holm correction applied. Second, post hoc analyses revealed faster VOTs for the distance 1 than 3 ($t(4420.41) = -3.18, p_{holm} < .01, estimate = -11.13, SE = 3.50$), yielding a classical priming distance effect (PDE). The PDE indicates that shorter distances between prime and the targets facilitated the naming of the targets, which is explained by the prime's semantic processing (see Figure 3).

Figure 3. *VOT (in ms) when presenting primes as digits*



Notes. Black lines illustrate VOTs when targets are named in French, red lines refer to German. The horizontal axis represents the prime-target distance 0, 1, 2, 3. Ribbons represent standard errors.

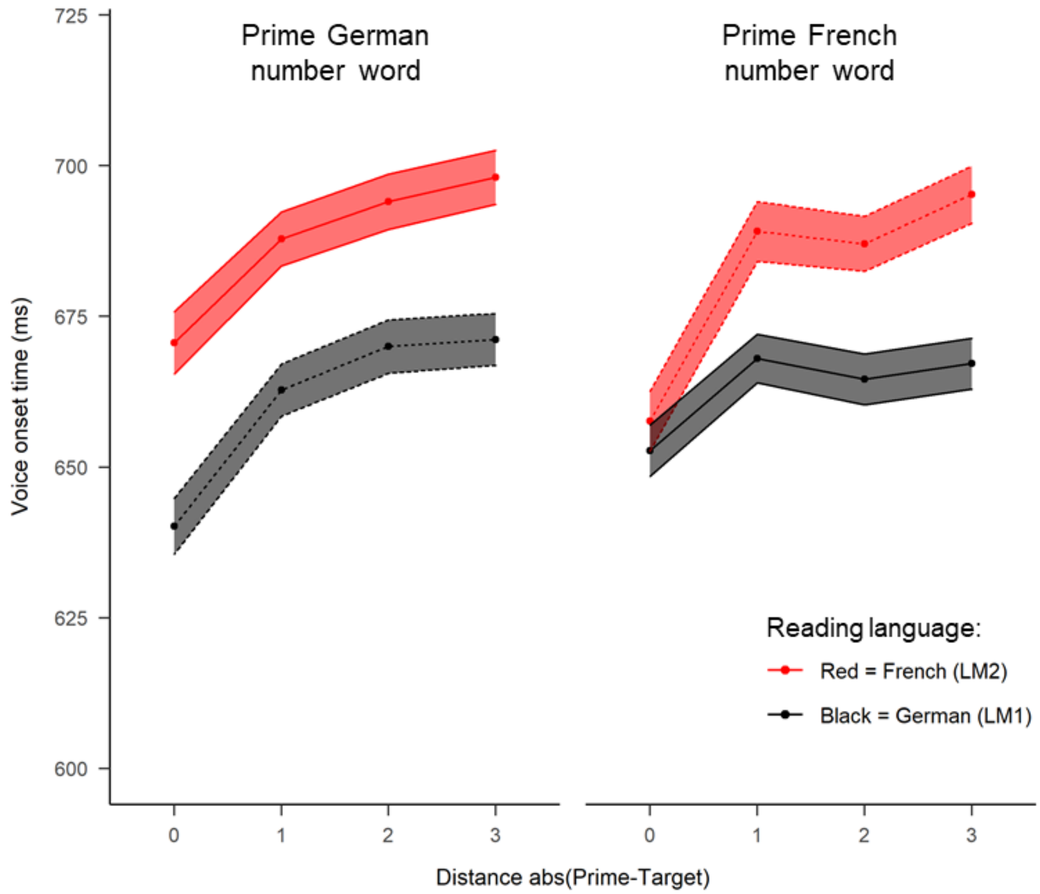
4.4.2 Number Words

The LMM in (A) was applied to number words, therefore including prime notation as a fixed factor with two levels (German number words, French number words). The LMM resulted in the main effect of the target naming language ($F(1,15.70) = 5.43, p < .05$), indicating a lexical cost for naming number words in French compared to German. Results also showed a main effect of distance ($F(3,8884.23) = 52.01, p < .001$). Furthermore, the two-way interaction between prime notation and target naming language was significant, ($F(1,8884.22) = 5.47, p < .05$). Critically, the three-way interaction between distance, prime notation, and target naming language was also significant, ($F(3,8884.21) = 3.31, p < .05$).

Posthoc decomposition of the three-way interaction was performed. The results show a repetition priming effect between all prime notations (number words in German and French) in combination with all target naming languages (German and French). That is, significant repetition priming (*i.e.* distance 0) occurred from German prime number words to targets to be named in German ($t(8884) = -4.43, p_{holm} < .001), estimate = -21.75\ , SE = 4.91$) and in French ($t(8884) = -3.60, p_{holm} < .001, estimate = -17.93,\ SE = 4.98$). Furthermore, repetition priming was also observed from French prime number words to targets named in German ($(8884) = -3.01, p_{holm} < .01, estimate = -14.74, SE = 4.90$) and in French ($t(8884) = -6.72, p_{holm} < .001, estimate = -33.13, SE = 4.93$), thus showing repetition priming from number words in German and in French, independently from the target naming language.

For the priming distance effect (PDE), we compared the distances 1 to 3 for each prime's notation and each target's naming language. These contrasts indicated a PDE for German number words, which was nearly significant when naming the targets in German ($t(8884) = -1.88, p_{holm} = .06, estimate = -9.21, SE = 4.90$) and significant, when naming in French ($t(8884) = -1.95, p_{holm} = .05, estimate = -9.68, SE = 4.96$). Critically, however in comparison, the number PDE effect was absent when French number words were used as primes, both when targets were named in German ($t(8884) = -0.03, p_{holm} = .97, estimate = -0.15, SE = 4.93$) and in French ($t(8884) = -1.14, p_{holm} = .25, estimate = -5.64, SE = 4.95$), see Figure 4. Hence, independently of the naming language, priming with German number words elicited a PDE, while priming with French number words did not. In other words, number words in French showed weak lexico-semantic access compared to number words in German.

Figure 4. *VOTs (in ms) when presenting German number word (left panel) and French number words (right panel) as primes.*



Notes. Black lines illustrate VOTs when targets are named in French, red lines refer to German. The horizontal axis represents the prime-target distance 0, 1, 2, 3. Ribbons represent standard errors. Segmented lines represent crossed language conditions (*i.e.* when the prime differs from the target naming language).

5 Discussion

The current study aimed to compare the lexical and semantic associations of single-digit numbers in bilinguals. To this aim, we tested highly proficient balanced German-French adult bilinguals who followed a school curriculum where the language for learning mathematics switches from German (LM1), *i.e.* 6 to 12 years old, to French (LM2) at 7th grade, *i.e.* 12 to 19 years old. Participants performed a number naming task in a semantic masked priming distance effect (PDE) design (Reynvoet et al. 2002) while their voice onset times (VOT) to the targets were measured. The target retrieval language (German vs French), the semantic distance between numbers (0,1,2,3), as well as the prime notations (no prime vs Digits vs French number words vs German number words) were manipulated within-subject. PDE was used as an estimate of the prime's semantic activation.

The overall results of the linear mixed models (LMM) analysis show the following pattern. *First*, the VOTs were slower when naming Indo-Arabic digits in French (LM2) than in German (LM1). This LM2 cost of about 20 ms is already significant in the no prime trials (*i.e.* "######") and overall in all prime notations. Since this general LM2 cost is not affected by the prime notation, we interpret it as arising from a lexical retrieval stage. *Second*, we found a PDE for Arabic digit primes: closer primes and targets elicited faster VOTs than distant pairs. On a methodological level, this PDE confirms the validity of the measures from the online platform (Kochari, 2019). Theoretically, since the PDE is found for both target naming languages, this result indicates that the lexico-semantic association activated by Indo-Arabic digits is language-independent, as predicted by the triple code model (Dehaene, 1992). *Third*, independently from the prime's notations, repetition priming trials (*i.e.* when the target and prime represent the same numbers, distance = 0) elicited faster VOTs than distance 1. We interpreted this result as strong lexico-lexical associations between Indo-Arabic digits and both of their verbal phonological German and French correspondents. Note that the repetition priming with Indo-Arabic digits might also be explained by low-level full visual overlap between prime and target, however, for number-word primes it must tap onto higher-level cognitive processes. Hence, an important interpretation from the repetition priming is that both LM1 and LM2 number words are associated at a lexical level with their exact Arabic digit match. The repetition priming worked for forward (*i.e.* LM1 prime number words to LM2 target naming language) and backward crossed prime-naming languages (*i.e.* LM2 to LM1), indicating a common process for both languages. *Fourth*, and critically, while trials primed with number words in German (LM1) resulted in a PDE, those with number word French (LM2) did not. These findings indicate weaker lexico-semantic associations from LM2 number words compared to LM1. Since LM2 number words were effective primes for the repetition priming but not for PDE, this suggests that LM2 number words are effectively processed at the lexical level, but have weak lexico-semantic associations with neighbouring numbers. In a nutshell, we thus observed a lexical retrieval cost for LM2 (French) when naming Arabic digit targets, and an additional cost in lexico-semantic access from LM2 number word primes, compared to LM1 (German).

5.1 Lexical retrieval cost

The lexical retrieval cost for naming Indo-Arabic digits in the LM2 replicates and extends previous findings. Similar LM2 costs were for example observed during arithmetic problem solving (Van Rinsveld et al., 2015) and two-digit number transcoding tasks (Lachelin et al., 2022) in a comparable sample that followed the Luxembourgish educational system. In general, those results align with psycholinguistics investigations indicating slower recognition and production for later learned words in object naming, word naming, and lexical decision tasks (i.e. Hirsh et al., 2003). Theoretically, the lexical LM2 cost might arise from language competition with the LM1 during the lexical retrieval stage. Since Indo-Arabic digits are language non-specific, they might have non-selective lexical access. Therefore, a digit possibly activates both languages' lexical correspondents. During this lexical competition, the LM2 cost would result from weaker lexical associations than present in LM1 number words. This lexical cost is predicted from multiple theories on bilingualism, such as the inhibition control hypothesis (Green, 1998), the revised hierarchical model (Kroll et al., 2010), and the bilingual activation model (BIA+) (Dijkstra & van Heuven, 2002). This prediction is also made by models specific to numerical cognition such as the bilingual encoding complex model (Bernardo, 2001), predicting stronger verbal codes in the language used for practicing arithmetic. Finally, regarding the proposed BTCM, it means that the weights of visual-verbal associations are weaker with the LM2 than with LM1 (see Figure 1, arrows A2 and A1, respectively).

Since number words are orthographically and phonologically longer in German than in French (see S1 Table 2), it is unlikely that this cost is due to the number words length effect (N. C. Ellis & Hennelly, 1980). In addition more transparent grapho-phonologically languages such as German have in general a more accentuated word length effect (Ziegler et al., 2001; Ziegler & Goswami, 2005). Hence the lexical cost observed for French (as LM2) compared to German (as LM1) might even be underestimated. Note that additionally, compared to monolinguals, this cost might add up to an already slower lexical retrieval in L1 for bilinguals (*i.e.* in picture naming tasks: Ivanova & Costa, 2008).

5.2 Lexico-semantic cost

We interpret the absence of PDE when priming with LM2-French number words compared to LM1 as elicited by an LM2 lexico-semantic cost. The cost appears at a later semantic association stage since priming LM2 number words elicited a repetition priming (*i.e.* "cinq" facilitated the naming of "5"), indicating LM2 number words were processed at an earlier lexical level. This cost is not appearing at target's lexical retrieval level since a PDE is observed with targets which are named in French and are being preceded by German (LM1) number word primes (*i.e.* "vier" facilitated the naming of " 5" as /cinq/). The presence of repetition priming in both languages but the absence of PDE selectively in the LM2 brings to the conclusion that number words prime's were identified in both languages speaking against a notation effect but rather to the strength of quantity activation (see Koechlin et al., 1999).

Our findings do not appear to align directly with a recent EEG study on bilingual arithmetic verification tasks in English and Spanish, which revealed a similar ERP (*i.e.* N400, marking semantic processing) for both languages (Cerda et al., 2019). However, the sample's language profile differed from the present study, as L2 was acquired very early (between 0 and 5 years) in comparison to the average 7 years of the present sample. Finally (Martinez-Lincoln et al., 2015) observed equivalent N400 peaks between mathematics performance in later and early learned languages when this was also the teacher's teaching language. This finding suggests the existence of late plasticity for arithmetic memory networks in specific cases, which might also exist for numbers. On the other side, weaker LM2 lexico-semantic associations fit with fMRI studies indicating more brain areas for solving arithmetics in the LM2 (Lin et al., 2019; Van Rinsveld et al., 2017; Wang et al., 2007). More extensive brain activation patterns are hence interpreted as more effortful and less efficient processes. Finally, the present study extends previous results concerning the PDE in bilinguals, such as Duyck et al., (2008) or Duyck and Brysbaert (2002) in that we found a cost with LM2 number word primes. Yet, it differs in that our experiment was designed to measure semantic mediation during number naming, rather than during translation. In addition, the task-relevant stimuli (*i.e.* primes) of the present study were Indo-Arabic digits, which are language-independent instead of language-dependent number words.

Weaker lexico-semantic associations for L2 fit with general psycholinguistic and specific numerical cognition models of bilingualism. For example, this prediction is made by the revised hierarchical model (Kroll et al., 2010), or the bilingual activation model (BIA+, Dijkstra & van Heuven, 2002) and the integrative multilink model (Dijkstra et al., 2019). Specific cognitive models for bilingual numerical cognition also predict weaker lexico-semantic association in the L2 (Duyck & Brysbaert, 2004). However, the bilingual encoding complex model does not predict weaker lexico-semantic association with the LM2, since it predicts an asymmetry in which the weaker lexical code systematically activates the stronger lexical code, while the present results indicate that the stronger lexical code (LM1) induces stronger semantic activations (Bernardo, 2001). Finally, regarding the proposed BTCM (see Figure 1), it would mean that in addition to the lexical association, the verbal lexico-semantic associations from number words are also weaker for LM2 than LM1 (arrows B2 and B1, respectively).

Possible sources of the LM2 cost

Cognitive models can provide an approach to explain the LM2 lexical and semantic cost. For example, connectionist models of bilingualism like the BIA+ (Dijkstra & van Heuven, 2002) and lately Multilink (Dijkstra et al., 2019) posit the existence of lexico-semantic nodes and connections which might vary in strengths. This theory fits the present study's result that number words would have weaker lexico-semantic connections in the LM2 than in the LM1. However, regarding the source of weaker lexico-semantic LM2 associations, we can only speculate. We suggest hence three potentially complementary accounts: age of acquisition (AoA), home language and bilingual word reading.

First, the general and mathematic specific (*i.e.* LM) language Age of Acquisition (AoA) is earlier for German than for French in our sample. The specific AoA of mathematical learning corresponds to the age at which mathematics is learned at school: from 1^{st} grade on in German (LM1) and from 7^{th} grade on in French (LM2). This corresponds to a strict definition of AoA, *i.e.* an "*intensive, systematic, and maintained exposure to his/her new language*" (Kovelman et al., 2008, p. 204), see also (A. W. Ellis & Lambon Ralph, 2000). Earlier AoA has neurocognitive effects and shapes the neuronal correlates of language and processes related to language. From a neuroscientific perspective, these differences are reflected in the recruitment of more brain regions when solving arithmetic in the L2 than the L1 (Martinez-Lincoln et al., 2015; Wang et al., 2007). Specifically, in a comparable sample more extensive brain activations for LM2 than LM1 arithmetic were also found (Van Rinsveld et al., 2017). A larger brain activation pattern could reflect less optimised cognitive networks when solving arithmetic in L2 or LM2. The weaker lexico-semantic LM2 association might therefore be explained by later AoA. Note, that language general vocabulary acquisition and math-specific vocabulary (*i.e.* number words) are confounded in the present design so that it is not possible to disentangle the AoA effect due to general versus math specific aspects. Nevertheless, since math education in LM2 lasts one year longer (*i.e.* 7 years) than in LM1 (*i.e.* 6 years), we can likely discard an exposure effect (or subjective frequency effect) related to number words in LM1.

Second, language proximity between LM1 and home language (HL) could also have played a role: 24 out of 32 participants reported Luxembourgish as the first most proficient language (5 as their second most proficient)[15]. Hence Luxembourgish was likely the HL of the present sample. Luxembourgish is linguistically close to German (linguistically as close as the German dialect Bayerisch is to German, see (Martini, 2021)), which might facilitate the acquisition of German compared to French. The stronger lexico-semantic associations of numbers in LM1 might therefore also have their source in the linguistic closeness between the HL (*i.e.* Luxembourgish) and LM1 (German). Note that the opposite is also possible: Luxembourgish might hinder French number word's lexico-semantic association. However, it must be noted that Luxembourgish is primarily an oral language; consequently German written number words are most likely acquired during school. Furthermore, the language of schooling (*i.e.* LM) is a stronger predictor of Arabic digit naming than HL, as is underlined from studies on bilinguals with Finnish HL and Swedish LM. A series of studies indicated faster Arabic digit naming in the Swedish LM than Finnish HL, already after three years of schooling (Chincotta & Underwood, 1997). Faster digit naming in the LM than in the HL was further accompanied by larger digit spans in LM (Chincotta & Underwood, 1996). Since schooling language seems be a stronger predictor than HL for Arabic digit naming, it might also be that linguistic proximity of LM1 (*i.e.* German) with the HL (*i.e.* Luxembourgish) might have facilitated LM1 lexico-semantic associations compared to (or even hindered) the LM2 association. Linguistic proximity is, however, probably not the only explaining factor since Luxembourgish is mainly a spoken language.

---

[15] The average AoA of for Luxembourgish was 3.8 years old (after excluding 13 reporting 0).

Third, the weak semantic activation by LM2 number words could also arise from reading-related differences as suggested by dual-route reading models (Coltheart et al., 1993). In this account, the LM1 number words would automatically and directly be associated with the semantics, while LM2 number word reading would rely on grapheme-phonological conversion mechanisms. This perspective could explain why LM2 number words elicited repetition priming as lexico-phonological facilitation but no PDE, given the short primes and stimulus onset asynchronies used in this study. However, it could not explain the repetition priming where French number words (*i.e.* SEPT) facilitated the reading of Indo-Arabic digits (*i.e.* 7) in German (/sieben/). The latter result indicates that LM2 number words benefit from a higher level of processing than grapheme-phonological conversion, *i.e.* lexical stage. Moreover, the present graduates from the multilingual Luxembourgish school system have high reading proficiencies in French, particularly for very frequent words, *i.e.* number words. With regards to reading, the language in which reading is first learned (a reading AoA effect), might alternatively explain the weaker lexico-semantic effect. Within this framework lexico-semantic associations in LM2 would be delayed rather than weaker. If this is the case, longer presentation times or stimulus onset times should result in a PDE in the LM2. Support for a role of reading proficiency comes from previous investigations indicating that Luxembourgish speakers' math performances are mediated by German reading comprehension (Greisen et al., 2021).

In conclusion, weaker LM2 lexico-semantic associations might originate from a combination of earlier AoA of LM1 than LM2 and the linguistic similarity between HL and LM1, as well as effects of reading. This would then foster stronger lexico-semantic association for LM1 number words than LM2. With regards to the proposed BTCM model, it means that the semantic associations with the language specific verbal codes depend on similar mechanisms than those at play for general language processes (*i.e.* AoA and language proximity).

5.3 Strengths, limitations, and perspectives

The present PDE task and the population recruited for our study have various strength and limitations regarding the type of stimuli used, their temporality and the language profiles of participants. Using numbers as stimuli has several advantages in comparison to general words and pictures, which are other stimuli typically used in bilingual investigations. For example, the same visual entry (*i.e* Indo-Arabic digits) is used for both languages' lexical accesses. Concerning semantics, numbers have very strong lexico-semantic overlap in different languages[16], in other words, they refer unambiguously to the same quantity and mathematical properties across languages. Finally, numbers are balanced for the frequency of exposure across languages (Dehaene & Mehler, 1992). However, the use of number words as primes comes with a limit regarding the interpretation of mechanisms related to arithmetic, since arithmetics are usually presented in Indo-Arabic digits (see Cerda et al., 2019).

The very short prime presentation might not have been long enough to reveal a potentially existing but delayed semantic association in LM2 compared to LM1. Indeed, while our results show that lexico-semantic associations cannot be elicited with short LM2 number word primes, the study is not informative about the exact temporality of these associations. Future studies should explore and compare the temporality of both languages' lexico-semantic access to see if the LM2 cost persists. Based on the observation that the PDE is symmetrical for small and large primes, previous PDE studies in monolinguals could exclude the role of counting in the observed response time pattern (Reynvoet et al., 2002). Future studies could nonetheless further explore the possibility that learning the counting sequence contributes to the observed effects in bilinguals. Also note that PDE designs are not suited for correlation, since they are typically observed as group rather than individual effects (see Sasanguie et al., 2011).

The language profiles were quite homogenous from the perspective of the language of learning mathematics and home language with 24 over 32 participants reported Luxembourgish as being their most proficient language. Nevertheless, the home language was not directly measured or controlled in this design, since Luxembourgish is linguistically close to German and prevalently oral (as for example Swiss-German). Also, the present multilingual sample reports to be proficient in 4.8 languages, this might have added additional concurrent verbal codes compared to a more exclusive bilingual sample. Notwithstanding the latter limitation, the strong homogeneity of the languages of math learning still makes the present multilingual participants a highly interesting and relevant population for the present research question. Moreover, comparing and understanding

[16] It might be argued that "un" in French is also a pronoun and might therefore be polysemantic.

the impact of languages of schooling has more practical implications, *i.e.* for school curriculum designs, than knowledge on the effect of home languages.

The present findings might have implication concerning multilingual education, since the LM2 cost observed here arises despite seven years of mathematics in French. Thus, switching language for math instruction in the context of such multilingual education curricula might eventually have detrimental impacts on learning mathematics. This becomes critical when considering that costs even increase for multiple digits transcoding (Lachelin et al., 2022) and are detrimentally affecting the speed of solving simple and complex arithmetic problems (Van Rinsveld et al., 2015). This could also indirectly be observable in that multilinguals prefer to use their more dominant language to solve arithmetic (Dewaele, 2007; Martini, 2021). Finally, it might also explain the use of lexical retrieval strategies in the LM1 in contrast to alternative strategies in the LM2 (i.e. visuospatial) as suggested by (Van Rinsveld et al., 2017). The cognitive cost entailed by sequentially bilingual math curricula should be considered given the hierarchical nature of math education and the strong implications for later individual achievement (Duncan et al., 2007). It might especially increase inequalities by hampering low achieving math students already struggling in a first language, as the second language is likely to add an additional difficulty towards their mathematical education,

Last, the present result of weaker LM2 lexical and semantic associations in proficient bilinguals might also be important regarding methodological aspects of numerical cognition research. For example, studies using number words might need to be cautious with including multilinguals in their samples, since including LM2 number words could affect the study outcomes.

5.4 Conclusion

Our results indicate that proficient bilinguals have two LM2 costs: one in lexical retrieval when naming Indo-Arabic digits and a second due to weaker lexico-semantic activation from LM2 number word primes. This cognitive component must be considered when switching a language of teaching and testing mathematical knowledge. The present results add up to previous studies revealing how bilingual school curricula involving a language switch might affect cognitive processes until adulthood. More generally, this study supports the importance of language in numerical cognition.

5.5 Constraints on Generality

We expect the result to generalize to high and low proficient bilinguals of other languages given a sequential acquisition of the languages and a strict control of the stimuli. For example both languages must have comparable lengths (or otherwise see Ellis & Hennelly, 1980). Given that the stimuli used here are slightly longer in German than French (see S1 Table 2), but the response times were shorter for the former, these should generalize to other stimuli. Note however that multi-digit numbers might present additional morpho-syntactic language differences which might explain additional cost or benefits when comparing languages (see Lachelin et al., 2022). Hence, we have no reason to believe that the results depend on other characteristics of the participants, materials, or context

6 Supplementary material 1:

S1 Table 1

List of all primes and targets of one condition.

Prime			Target	Distance
Indo-Arabic digit	German Number Words	French Number Words		
1	EINS	UN	1	0
2	ZWEI	DEUX	2	0
3	DREI	TROIS	3	0
4	VIER	QUATRE	4	0
5	FÜNF	CINQ	5	0
6	SECHS	SIX	6	0
7	SIEBEN	SEPT	7	0
8	ACHT	HUIT	8	0
9	NEUN	NEUF	9	0
1	EINS	UN	2	1
2	ZWEI	DEUX	1	1
2	ZWEI	DEUX	1	1
2	ZWEI	DEUX	3	1
3	DREI	TROIS	2	1
3	DREI	TROIS	4	1
4	VIER	QUATRE	3	1
4	VIER	QUATRE	5	1
5	FÜNF	CINQ	4	1
5	FÜNF	CINQ	6	1
6	SECHS	SIX	5	1
6	SECHS	SIX	7	1
7	SIEBEN	SEPT	6	1

7	SIEBEN	SEPT	8	1
8	ACHT	HUIT	7	1
8	ACHT	HUIT	9	1
8	ACHT	HUIT	9	1
9	NEUN	NEUF	8	1
1	EINS	UN	3	2
2	ZWEI	DEUX	4	2
3	DREI	TROIS	1	2
3	DREI	TROIS	1	2
3	DREI	TROIS	5	2
4	VIER	QUATRE	2	2
4	VIER	QUATRE	2	2
4	VIER	QUATRE	6	2
5	FÜNF	CINQ	3	2
5	FÜNF	CINQ	7	2
6	SECHS	SIX	4	2
6	SECHS	SIX	8	2
6	SECHS	SIX	8	2
7	SIEBEN	SEPT	5	2
7	SIEBEN	SEPT	9	2
7	SIEBEN	SEPT	9	2
8	ACHT	HUIT	6	2
9	NEUN	NEUF	7	2
1	EINS	UN	4	3
2	ZWEI	DEUX	5	3
3	DREI	TROIS	6	3
4	VIER	QUATRE	1	3
4	VIER	QUATRE	1	3
4	VIER	QUATRE	7	3
4	VIER	QUATRE	7	3
5	FÜNF	CINQ	2	3
5	FÜNF	CINQ	2	3

5	FÜNF	CINQ	8	3
5	FÜNF	CINQ	8	3
6	SECHS	SIX	3	3
6	SECHS	SIX	3	3
6	SECHS	SIX	9	3
6	SECHS	SIX	9	3
7	SIEBEN	SEPT	4	3
8	ACHT	HUIT	5	3
9	NEUN	NEUF	6	3
1	EINS	UN	1	0
2	ZWEI	DEUX	2	0
3	DREI	TROIS	3	0
4	VIER	QUATRE	4	0
5	FÜNF	CINQ	5	0
6	SECHS	SIX	6	0
7	SIEBEN	SEPT	7	0
8	ACHT	HUIT	8	0
9	NEUN	NEUF	9	0
######			1	-Filler-
######			2	-Filler-
######			3	-Filler-
######			4	-Filler-
######			5	-Filler-
######			6	-Filler-
######			7	-Filler-
######			8	-Filler-
######			9	-Filler-
######			1	-Filler-
######			2	-Filler-
######			3	-Filler-
######			4	-Filler-
######			5	-Filler-

######	6	-Filler-
######	7	-Filler-
######	8	-Filler-
######	9	-Filler-

Note: design of the 234 trials corresponding to one language condition with the following primes: 72 Indo-Arabic Digits, 72 number words in German and 72 number words in French and 18 filler trials (234 trials per condition).

S1 Table 2

German and French number words' linguistic characteristics comparison

Digit	German				French			
		Length (Ortho)	Length (Phono)	Frequency (per-million)		Length (Ortho)	Length (Phono)	Frequency (per-million)
1	EINS	4	3	257.57	UN	2	1	12685.55
2	ZWEI	4	3	994.37	DEUX	4	2	1009.22
3	DREI	4	3	453.48	TROIS	5	4	384.48
4	VIER	4	3	210.95	QUATRE	6	4	152.16
5	FÜNF	4	4	177.64	CINQ	4	3	163.06
6	SECHS	5	4	114.30	SIX	3	3	118.19
7	SIEBEN	6	5	72.09	SEPT	4	3	67.38
8	ACHT	4	3	141.30	HUIT	4	3	59.41
9	NEUN	4	3	56.77	NEUF	4	3	41.25
Mean:		4.3	3.4	275.4		4	2.9	1631.2

Note. Comparison of number words in French and German: French number words are on average shorter both phonologically and orthographically. The relative frequency of French number words is higher/equivalent than in German. However, in German "Eins" refers exclusively to a number word, while "ein" the equivalent of the adjective "a" in English, is more frequent (11034.61 /million) with a similar frequency as "un" in French. Retrieved from https://clearpond.northwestern.edu/.

7 References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*, *128*(3), 466–478. https://doi.org/10.1016/j.actpsy.2008.03.014

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Bahnmueller, J., Moeller, K., Mann, A., & Nuerk, H.-C. (2015). On the limits of language influences on numerical cognition – no inversion effects in three-digit number magnitude processing in adults. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01216

Bahnmueller, J., Nuerk, H.-C., & Moeller, K. (2018). A Taxonomy Proposal for Types of Interactions of Language and Place-Value Processing in Multi-Digit Numbers. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01024

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00328

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004a). ADAPT: A Developmental, Asemantic, and Procedural Model for Transcoding From Verbal to Arabic Numerals. *Psychological Review*, *111*(2), 368–394. https://doi.org/10.1037/0033-295X.111.2.368

Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004b). ADAPT: A Developmental, Asemantic, and Procedural Model for Transcoding From Verbal to Arabic Numerals. *Psychological Review*, *111*(2), 368–394. https://doi.org/10.1037/0033-295X.111.2.368

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*(3), 201–221. https://doi.org/10.1016/S0010-0277(02)00178-6

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bernardo, A. B. I. (2001). Asymmetric activation of number codes in bilinguals: Further evidence for the encoding complex model of number processing. *Memory & Cognition*, *29*(7), 968–976. https://doi.org/10.3758/BF03195759

Bernardo, A. B. I. (2005). Language and Modeling Word Problems in Mathematics Among Bilinguals. *The Journal of Psychology*, *139*(5), 413–425. https://doi.org/10.3200/JRLP.139.5.413-425

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General*, *124*(4), 434–452. https://doi.org/10.1037/0096-3445.124.4.434

Bürkner, P.-C. (2017). **brms**: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Bylund, E., Antfolk, J., Abrahamsson, N., Olstad, A. M. H., Norrman, G., & Lehtonen, M. (2023). Does bilingualism come with linguistic costs? A meta-analytic review of the bilingual lexical deficit. *Psychonomic Bulletin & Review*, *30*(3), 897–913. https://doi.org/10.3758/s13423-022-02136-7

Camos, V. (2008). Low working memory capacity impedes both efficiency and learning of number transcoding in children. *Journal of Experimental Child Psychology*, *99*(1), 37–57. https://doi.org/10.1016/j.jecp.2007.06.006

Campbell, J. (1995). Mechanisms of Simple Addition and Multiplication: A Modified Network-interference Theory and Simulation. *Mathematical Cognition*, *1*, 121–164.

Campbell, J. I. D., & Epp, L. J. (2004). An Encoding-Complex Approach to Numerical Cognition in Chinese-English Bilinguals. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *58*(4), 229–244. https://doi.org/10.1037/h0087447

Cerda, V. R., Grenier, A. E., & Wicha, N. Y. Y. (2019). Bilingual children access multiplication facts from semantic memory equivalently across languages: Evidence from the N400. *Brain and Language*, *198*, 104679. https://doi.org/10.1016/j.bandl.2019.104679

Chincotta, D., & Underwood, G. (1996). Mother tongue, language of schooling and bilingual digit span. *British Journal of Psychology*, *87*(2), 193–208. https://doi.org/10.1111/j.2044-8295.1996.tb02585.x

Chincotta, D., & Underwood, G. (1997). Speech Rate Estimates, Language of Schooling and Bilingual Digit Span. *European Journal of Cognitive Psychology*, *9*(3), 325–348. https://doi.org/10.1080/713752562

Chrisomalis, S. (2010). *Numerical Notation: A Comparative History*. Cambridge University Press.

Clayton, F. J., Copper, C., Steiner, A. F., Banfi, C., Finke, S., Landerl, K., & Göbel, S. M. (2020). Two-digit number writing and arithmetic in Year 1 children: Does number word inversion matter? *Cognitive Development*, *56*, 100967. https://doi.org/10.1016/j.cogdev.2020.100967

Cohen, J., Mac Whinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers*, *25*(2), 257–271.

Colomé, àngels, Laka, I., & Sebastián-Gallés, N. (2010). Language effects in addition: How you say it counts. *Quarterly Journal of Experimental Psychology*, *63*(5), 965–983. https://doi.org/10.1080/17470210903134377

Colomé, À. (2001). Lexical Activation in Bilinguals' Speech Production: Language-Specific or Language-Independent? *Journal of Memory and Language*, *45*(4), 721–736. https://doi.org/10.1006/jmla.2001.2793

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589–608. https://doi.org/10.1037/0033-295X.100.4.589

Comrie, B. (2013). Numeral bases (v2020.3). In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Zenodo. https://doi.org/10.5281/zenodo.7385533

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002

de Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. Psychology Press. https://doi.org/10.4324/9780203841228

De Visscher, A., & Noël, M.-P. (2014). The detrimental effect of interference in multiplication facts storing: Typical development and individual differences. *Journal of Experimental Psychology: General*, *143*(6), 2380–2400. https://doi.org/10.1037/xge0000029

De Vos, T. (1992). Tempo test rekenen (TTR)[Arithmetic number fact test]. *Nijmegen: Berkhout*.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*(1), 1–29. https://doi.org/10.1016/0010-0277(92)90030-l

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence. *Science*, *284*(5416), 970–974. https://doi.org/10.1126/science.284.5416.970

Del Maschio, N., & Abutalebi, J. (2019). Language Organization in the Bilingual and Multilingual Brain. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the Neuroscience of Multilingualism* (pp. 197–213). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119387725.ch9

Deloche, G., & Seron, X. (1982). From three to 3: A differential analysis of skills in transcoding quantities between patients with Broca's and Wernicke's aphasia. *Brain*, *105*(4), 719–733. https://doi.org/10.1093/brain/105.4.719

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, *82*(1), 64–81. https://doi.org/10.1348/2044-8279.002002

Dewaele, J.-M. (2007). Multilinguals' language choice for mental calculation. *Intercultural Pragmatics*, *4*(3). https://doi.org/10.1515/IP.2007.017

Dijkstra, T. (2005). Bilingual Visual Word Recognition and Lexical Access. In *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). Oxford University Press. https://doi.org/10.1017/S0272263107210071

Dijkstra, T., & Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Wahl, A., Buytenhuijs, F., Halem, N. V., Al-Jibouri, Z., Korte, M. D., & Rekké, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Dotan, D., & Friedmann, N. (2018). A cognitive model for multidigit number reading: Inferences from individuals with selective impairments. *Cortex*, *101*, 249–281. https://doi.org/10.1016/j.cortex.2017.10.025

Dowker, A., & Nuerk, H.-C. (2016). Editorial: Linguistic Influences on Mathematics. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01035

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Duyck, W., & Brysbaert, M. (2002). What number translation studies can teach us about the lexico-semantic organisation in bilinguals. *Psychologica Belgica*, *42*(3), 151–175.

Duyck, W., & Brysbaert, M. (2004). Forward and Backward Number Translation Requires Conceptual Mediation in Both Balanced and Unbalanced Bilinguals. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 889–906. https://doi.org/10.1037/0096-1523.30.5.889

Duyck, W., Depestel, I., Fias, W., & Reynvoet, B. (2008). Cross-lingual numerical distance priming with second-language number words in native- to third-language number word translation. *Quarterly Journal of Experimental Psychology*, *61*(9), 1281–1290. https://doi.org/10.1080/17470210802000679

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of Acquisition Effects in Adult Lexical Processing Reflect Loss of Plasticity in Maturing Systems: Insights From Connectionist Networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1103–1123. https://doi.org/10.1037/0278-7393.26.5.1103

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology*, *71*(1), 43–51. https://doi.org/10.1111/j.2044-8295.1980.tb02728.x

Fiebach, C. J., Friederici, A. D., Müller, K., von Cramon, D. Y., & Hernandez, A. E. (2003). Distinct brain representations for early and late learned words. *NeuroImage*, *19*(4), 1627–1637. https://doi.org/10.1016/S1053-8119(03)00227-1

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: A unified JavaScript framework for online studies. *International Conference on Computational Social Science (Cologne)*. https://www.labvanced.com/

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*(3), 819–824. https://doi.org/10.1016/j.cognition.2008.04.007

Frenck-Mestre, C., & Vaid, J. (1993). Activation of number facts in bilinguals. *Memory & Cognition*, *21*(6), 809–818. https://doi.org/10.3758/BF03202748

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/10.1016/j.edurev.2013.05.003

Garcia, O., Faghihi, N., Raola, A. R., & Vaid, J. (2021). Factors influencing bilinguals’ speed and accuracy of number judgments across languages: A meta-analytic review. *Journal of Memory and Language*, *118*, 104211. https://doi.org/10.1016/j.jml.2020.104211

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, *8*(3), 259. https://doi.org/10.1007/BF03174081

Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of Arithmetical Competencies in Chinese and American Children: Influence of Age, Language, and Schooling. *Child Development*, *67*(5), 2022–2044. https://doi.org/10.1111/j.1467-8624.1996.tb01841.x

Göbel, S. M., Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H.-C. (2014). Language affects symbolic arithmetic in children: The case of number word inversion. *Journal of Experimental Child Psychology*, *119*, 17–25. https://doi.org/10.1016/j.jecp.2013.10.001

Göbel, S. M., Watson, S. E., Lervåg, A., & Hulme, C. (2014). Children's Arithmetic Development: It Is Number Knowledge, Not the Approximate Number Sense, That Counts. *Psychological Science*, *25*(3), 789–798. https://doi.org/10.1177/0956797613516471

Grabner, R. H., Saalbach, H., & Eckstein, D. (2012a). Language-Switching Costs in Bilingual Mathematics Learning. *Mind, Brain, and Education*, *6*(3), 147–155. https://doi.org/10.1111/j.1751-228X.2012.01150.x

Grabner, R. H., Saalbach, H., & Eckstein, D. (2012b). Language-Switching Costs in Bilingual Mathematics Learning. *Mind, Brain, and Education*, *6*(3), 147–155. https://doi.org/10.1111/j.1751-228X.2012.01150.x

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

Greisen, M., Georges, C., Hornung, C., Sonnleitner, P., & Schiltz, C. (2021). Learning mathematics with shackles: How lower reading comprehension in the language of mathematics instruction accounts for lower mathematics achievement in speakers of different home languages. *Acta Psychologica*, *221*, 103456. https://doi.org/10.1016/j.actpsy.2021.103456

Grosjean, F. (2001). *The Bilingual's Language Modes*. Blackwell Publishing.

Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.

Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.

Hahn, C. G. K., Saalbach, H., & Grabner, R. H. (2017). Language-dependent knowledge acquisition: Investigating bilingual arithmetic learning. *Bilingualism: Language and Cognition*, *22*(1), 47–57. https://doi.org/10.1017/S1366728917000530

Hahn, C. G. K., Saalbach, H., & Grabner, R. H. (2019). Language-dependent knowledge acquisition: Investigating bilingual arithmetic learning. *Bilingualism: Language and Cognition*, *22*(1), 47–57. https://doi.org/10.1017/S1366728917000530

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). *Individual differences in non-verbal number acuity correlate with maths achievement*. *455*, 655–668. https://doi.org/10.1038/nature07246

Haspelmath, M., Dryer, M. S., Gil, D., & Comrie, B. (2005). *The World Atlas of Language Structures*. Oxford Univ. Press.

Hernandez, A. E. (2013). *The bilingual brain*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199828111.001.0001

Hirsh, K. W., Morrison, C. M., Gaset, S., & Carnicer, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition*, *6*(2), 117–128. https://doi.org/10.1017/S136672890300107X

Ifrah, G., & Bellos, D. (2000). *The universal history of numbers: From prehistory to the invention of the computer*. Wiley.

Imbo, I., Vanden Bulcke, C., De Brauwer, J., & Fias, W. (2014). Sixty-four or four-and-sixty? The influence of language and working memory on children's number transcoding. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00313

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277–288. https://doi.org/10.1016/j.actpsy.2007.06.003

Kempert, S., Saalbach, H., & Hardy, I. (2011). Cognitive benefits and costs of bilingualism in elementary school students: The case of mathematical word problems. *Journal of Educational Psychology*, *103*(3), 547–561. https://doi.org/10.1037/a0023619

Klaus, J., & Schriefers, H. (2019). Bilingual Word Production. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the Neuroscience of Multilingualism* (pp. 214–229). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119387725.ch10

Kochari, A. R. (2019). Conducting Web-Based Experiments for Numerical Cognition Research. *Journal of Cognition*, *2*(1), 39. https://doi.org/10.5334/joc.85

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed Numbers: Exploring the Modularity of Numerical Representations With Masked and Unmasked Semantic Priming. *Journal of Experimental Psychology: Human Perception and Performance*, 24. https://doi.org/10.1037/0096-1523.25.6.1882

Kovelman, I., Baker, S. A., & Petitto, L.-A. (2008). Age of first bilingual language exposure as a new window into bilingual reading development. *Bilingualism: Language and Cognition*, *11*(2), 203–223. https://doi.org/10.1017/S1366728908003386

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01795

Krinzinger, H., Gregoire, J., Desoete, A., Kaufmann, L., Nuerk, H.-C., & Willmes, K. (2011). Differential Language Effects on Numerical Skills in Second Grade. *Journal of Cross-Cultural Psychology*, *42*(4), 614–629. https://doi.org/10.1177/0022022111406252

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, *13*(3), 373–381. https://doi.org/10.1017/S136672891000009X

Lachelin, R., Marinova, M., Reynvoet, B., & Schiltz, C. (2023). Weaker semantic priming effects with number words in the second language of math learning. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001526

Lachelin, R., Rinsveld, A. van, Poncin, A., & Schiltz, C. (2022). Number transcoding in bilinguals—A transversal developmental study. *PLOS ONE*, *17*(8), e0273391. https://doi.org/10.1371/journal.pone.0273391

Lê, M.-L., & Noël, M.-P. (2021). Preschoolers' mastery of advanced counting: The best predictor of addition skills 2 years later. *Journal of Experimental Child Psychology*, *212*, 105252. https://doi.org/10.1016/j.jecp.2021.105252

Lê, M.-L. T., & Noël, M.-P. (2020). Transparent number-naming system gives only limited advantage for preschooler's numerical development: Comparisons of Vietnamese and French-speaking children. *PLOS ONE*, *15*(12), e0243472. https://doi.org/10.1371/journal.pone.0243472

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia*, *41*(14), 1942–1958. https://doi.org/10.1016/S0028-3932(03)00123-4

Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means* [Manual]. https://CRAN.R-project.org/package=emmeans

Lin, J.-F. L., Imada, T., & Kuhl, P. K. (2012). Mental Addition in Bilinguals: An fMRI Study of Task-Related and Performance-Related Activation. *Cerebral Cortex*, *22*(8), 1851–1861. https://doi.org/10.1093/cercor/bhr263

Lin, J.-F. L., Imada, T., & Kuhl, P. K. (2019). Neuroplasticity, bilingualism, and mental mathematics: A behavior-MEG study. *Brain and Cognition*, *134*, 122–134. https://doi.org/10.1016/j.bandc.2019.03.006

Lonnemann, J., & Yan, S. (2015). Does number word inversion affect arithmetic processes in adults? *Trends in Neuroscience and Education*, *4*(1), 1–5. https://doi.org/10.1016/j.tine.2015.01.002

Major, C. S., Paul, J. M., & Reeve, R. A. (2017). TEMA and Dot Enumeration Profiles Predict Mental Addition Problem Solving Speed Longitudinally. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02263

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115. https://doi.org/10.1017/S1366728903001068

Marinova, M., Georges, C., Guillaume, M., Reynvoet, B., Schiltz, C., & Van Rinsveld, A. (2021). Automatic integration of numerical formats examined with frequency-tagged EEG. *Scientific Reports*, *11*(1), 21405. https://doi.org/10.1038/s41598-021-00738-0

Marsh, L. G., & Maki, R. H. (1976). Efficiency of arithmetic operations in bilinguals as a function of language. *Memory & Cognition*, *4*(4), 459–464. https://doi.org/10.3758/BF03213203

Martinez-Lincoln, A., Cortinas, C., & Wicha, N. Y. Y. (2015). Arithmetic memory networks established in childhood are changed by experience in adulthood. *Neuroscience Letters*, *0*, 325–330. https://doi.org/10.1016/j.neulet.2014.11.010

Martini, S. (2021). *The influence of language on mathematics in a multilingual educational setting*. University of Luxembourg.

McClain, L., & Huang, J. Y. S. (1982). Speed of simple arithmetic in bilinguals. *Memory & Cognition*, *10*(6), 591–596. https://doi.org/10.3758/BF03202441

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, *44*(1–2), 107–157. https://doi.org/10.1016/0010-0277(92)90052-J

McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition*, *4*(2), 171–196. https://doi.org/10.1016/0278-2626(85)90069-7

McClung, N. A., & Arya, D. J. (2018). Individual Differences in Fourth-Grade Math Achievement in Chinese and English. *Frontiers in Education*, *3*, 29. https://doi.org/10.3389/feduc.2018.00029

Meeuwissen, M., Roelofs, A., & Levelt, W. J. M. (2003). Planning levels in naming and reading complex numerals. *Memory & Cognition*, *31*(8), 1238–1248. https://doi.org/10.3758/BF03195807

Miller, K. F., Smith, C. M., Zhu, J., & Zhang, H. (1995). Preschool Origins of Cross-National Differences in Mathematical Competence: The Role of Number-Naming Systems. *Psychological Science*, *6*(1), 56–60.

Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development*, *2*(3), 279–305. https://doi.org/10.1016/S0885-2014(87)90091-8

Ministère de l'Éducation Nationale. (2022, April 20). Languages in Luxembourg Schools. https://men.public.lu/en/themes-transversaux/langues-ecole-luxembourgeoise.html

Miura, I. T., Kim, C. C., Chang, C.-M., & Okamoto, Y. (1988). Effects of Language Characteristics on Children's Cognitive Representation of Number: Cross-National Comparisons. *Child Development*, *59*(6), 1445. https://doi.org/10.2307/1130659

Moeller, K., Shaki, S., Göbel, S. M., & Nuerk, H.-C. (2015). Language influences number processing – A quadrilingual study. *Cognition*, *136*, 150–155. https://doi.org/10.1016/j.cognition.2014.11.003

Moeller, K., Zuber, J., Olsen, N., Nuerk, H.-C., & Willmes, K. (2015). Intransparent German number words complicate transcoding – a translingual comparison with Japanese. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00740

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), Article 5109. https://doi.org/10.1038/2151519a0

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious Masked Priming Depends on Temporal Attention. *Psychological Science*, *13*(5), 416–424. https://doi.org/10.1111/1467-9280.00474

Naccache, L., & Dehaene, S. (2001). The Priming Method: Imaging Unconscious Repetition Priming Reveals an Abstract Representation of Number in the Parietal Lobes. *Cerebral Cortex*, *11*(10), 966–974. https://doi.org/10.1093/cercor/11.10.966

Negen, J., & Sarnecka, B. W. (2009). *Young children's number-word knowledge predicts their performance on a nonlinguistic number task*. https://escholarship.org/uc/item/1q03q75z

Negen, J., & Sarnecka, B. W. (2012). Number-Concept Acquisition and General Vocabulary Development. *Child Development*, *83*(6), 2019–2027. https://doi.org/10.1111/j.1467-8624.2012.01815.x

Notebaert, K., Pesenti, M., & Reynvoet, B. (2010). The neural origin of the priming distance effect: Distance-dependent recovery of parietal activation using symbolic magnitudes. *Human Brain Mapping*, *31*(5), 669–677. https://doi.org/10.1002/hbm.20896

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition*, *111*(2), 275–279. https://doi.org/10.1016/j.cognition.2009.02.002

Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition*, *82*(1), B25–B33. https://doi.org/10.1016/S0010-0277(01)00142-1

Nuerk, H.-C., Weger, U., & Willmes, K. (2004). On the Perceptual Generality of the Unit-Decade Compatibility Effect. *Experimental Psychology*, *51*(1), 72–79. https://doi.org/10.1027/1618-3169.51.1.72

Nuerk, H.-C., Weger, U., & Willmes, K. (2005). Language effects in magnitude comparison: Small, but not irrelevant. *Brain and Language*, *92*(3), 262–277. https://doi.org/10.1016/j.bandl.2004.06.107

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, *15*(2), 202–206. https://doi.org/10.1016/j.conb.2005.03.007

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503. https://doi.org/10.1126/science.1102085

Pitt, B., Gibson, E., & Piantadosi, S. T. (2022). *Exact number concepts are limited to the verbal count range*. *33*(3), 371–381. https://doi.org/10.1177/09567976211034502

Pixner, S., Zuber, J., Heřmanová, V., Kaufmann, L., Nuerk, H.-C., & Moeller, K. (2011). One language, two number-word systems and many problems: Numerical cognition in the Czech language. *Research in Developmental Disabilities*, *32*(6), 2683–2689. https://doi.org/10.1016/j.ridd.2011.06.004

Poncin, A., Rinsveld, A. V., & Schiltz, C. (2020). *Units first or tens first: How bilingualism affects two-digit number transcoding?* PsyArXiv. https://doi.org/10.31234/osf.io/sg7ea

Poncin, A., Van Rinsveld, A., & Schiltz, C. (2019). Units-first or tens-first: Does language matter when processing visually presented two-digit numbers? *Quarterly Journal of Experimental Psychology*, *73*(5), 726–738. https://doi.org/10.1177/1747021819892165

Power, R. J. D., & Dal Martello, M. F. (1990). The dictation of Italian numerals. *Language and Cognitive Processes*, *5*(3), 237–254. https://doi.org/10.1080/01690969008402106

Prior, A., Katz, M., Mahajna, I., & Rubinsten, O. (2015). Number word structure in first and second language influences arithmetic skills. *Frontiers in Psychology*, *6*(MAR), 266. https://doi.org/10.3389/fpsyg.2015.00266

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859–862. https://doi.org/10.3758/BF03192979

R Core Team. (2013). *R: A language and environment for statistical computing*. http://www.R-project.org/

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, *72*(2), 191–201. https://doi.org/10.1016/S0010-0277(99)00048-7

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic priming in number naming. *The Quarterly Journal of Experimental Psychology Section A*, *55*(4), 1127–1139. https://doi.org/10.1080/02724980244000116

Reynvoet, B., De Smedt, B., & Van den Bussche, E. (2009). Children's representation of symbolic magnitude: The development of the priming distance effect. *Journal of Experimental Child Psychology*, *103*(4), 480–489. https://doi.org/10.1016/j.jecp.2009.01.007

Rodic, M., Zhou, X., Tikhomirova, T., Wei, W., Malykh, S., Ismatulina, V., Sabirova, E., Davidova, Y., Tosto, M. G., Lemelin, J.-P., & Kovas, Y. (2015). Cross-cultural investigation into cognitive underpinnings of individual differences in early arithmetic. *Developmental Science*, *18*(1), 165–174. https://doi.org/10.1111/desc.12204

RStudio Team. (2020). *RStudio: Integrated development environment for r* [Manual]. http://www.rstudio.com/

Saalbach, H., Eckstein, D., Andri, N., Hobi, R., & Grabner, R. H. (2013). When language of instruction and language of application differ: Cognitive costs of bilingual mathematics learning. *Learning and Instruction*, *26*, 36–44. https://doi.org/10.1016/j.learninstruc.2013.01.002

Salillas, E., Barraza, P., & Carreiras, M. (2015). Oscillatory Brain Activity Reveals Linguistic Prints in the Quantity Code. *PLOS ONE*, *10*(4), e0121434. https://doi.org/10.1371/journal.pone.0121434

Salillas, E., & Carreiras, M. (2014). Core number representations are shaped by language. *Cortex*, *52*, 1–11. https://doi.org/10.1016/j.cortex.2013.12.009

Salillas, E., & Martínez, A. (2018). Linguistic Traces in Core Numerical Knowledge: An Approach From Bilingualism. In *Language and Culture in Mathematical Cognition* (pp. 173–196). Elsevier. https://linkinghub.elsevier.com/retrieve/pii/B9780128125748000080

Salillas, E., & Wicha, N. Y. Y. (2012). Early Learning Shapes the Memory Networks for Arithmetic: Evidence From Brain Potentials in Bilinguals. *Psychological Science*, *23*(7), 745–755. https://doi.org/10.1177/0956797612446347

Sasanguie, D., Defever, E., Van den Bussche, E., & Reynvoet, B. (2011). The reliability of and the relation between non-symbolic numerical distance effects in comparison, same-different judgments and priming. *Acta Psychologica*, *136*(1), 73–80. https://doi.org/10.1016/j.actpsy.2010.10.004

Sasanguie, D., & Reynvoet, B. (2014). Adults' Arithmetic Builds on Fast and Automatic Processing of Arabic Digits: Evidence from an Audiovisual Matching Paradigm. *PLOS ONE*, *9*(2), e87739. https://doi.org/10.1371/journal.pone.0087739

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., & Smedt, B. D. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3), e12372. https://doi.org/10.1111/desc.12372

Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology*, *12*(3), 281–300. https://doi.org/10.1111/j.2044-835X.1994.tb00635.x

Singmann, H. (2021). *Mixed Model Reanalysis of RT data* [Computer software]. https://cran.r-project.org/web/packages/afex/vignettes/afex_mixed_example.html

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *afex: Analysis of factorial experiments* [Computer software]. https://CRAN.R-project.org/package=afex

Spaepen, E., Coppola, M., Flaherty, M., Spelke, E., & Goldin-Meadow, S. (2013). Generating a lexicon without a language model: Do words for number count? *Journal of Memory and Language*, *69*(4), 496–505. https://doi.org/10.1016/j.jml.2013.05.004

Spelke, E. S., & Tsivkin, S. (2001a). Initial knowledge and conceptual change: Space and number. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (pp. 70–98). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.005

Spelke, E. S., & Tsivkin, S. (2001b). Language and number: A bilingual training study. *Cognition*, *78*(1), 45–88. https://doi.org/10.1016/S0010-0277(00)00108-6

Steiner, A. F., Banfi, C., Finke, S., Kemény, F., Clayton, F. J., Göbel, S. M., & Landerl, K. (2021). Twenty-four or four-and-twenty: Language modulates cross-modal matching for multidigit numbers in children and adults. *Journal of Experimental Child Psychology*, *202*, 104970. https://doi.org/10.1016/j.jecp.2020.104970

Steiner, A. F., Finke, S., Clayton, F. J., Banfi, C., Kemény, F., Göbel, S. M., & Landerl, K. (2021). Language effects in early development of number writing and reading. *Journal of Numerical Cognition*, *7*(3), 368–387. https://doi.org/10.5964/jnc.6929

Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual Word Recognition in a Sentence Context. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00174

van der Ven, S. H. G., Klaiber, J. D., & van der Maas, H. L. J. (2017). Four and twenty blackbirds: How transcoding ability mediates the relationship between visuospatial working memory and math in a language with inversion. *Educational Psychology*, *37*(4), 487–505. https://doi.org/10.1080/01443410.2016.1150421

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

van Marle, K., Chu, F. W., Li, Y., & Geary, D. C. (2014). Acuity of the approximate number system and preschoolers' quantitative development. *Developmental Science*, *17*(4), 492–505. https://doi.org/10.1111/desc.12143

Van Rinsveld, A., Brunner, M., Landerl, K., Schiltz, C., & Ugen, S. (2015). The relation between language and arithmetic in bilinguals: Insights from different stages of language acquisition. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00265

Van Rinsveld, A., Dricot, L., Guillaume, M., Rossion, B., & Schiltz, C. (2017). Mental arithmetic in the bilingual brain: Language matters. *Neuropsychologia*, *101*, 17–29. https://doi.org/10.1016/j.neuropsychologia.2017.05.009

Van Rinsveld, A., & Schiltz, C. (2016). Sixty-twelve = Seventy-two? A cross-linguistic comparison of children's number transcoding. *British Journal of Developmental Psychology*, *34*(3), 461–468. https://doi.org/10.1111/bjdp.12151

Van Rinsveld, A., Schiltz, C., Brunner, M., Landerl, K., & Ugen, S. (2016). Solving arithmetic problems in first and second language: Does the language context matter? *Learning and Instruction*, *42*, 72–82. https://doi.org/10.1016/j.learninstruc.2016.01.003

Van Rinsveld, A., Schiltz, C., Landerl, K., Brunner, M., & Ugen, S. (2016). Speaking two languages with different number naming systems: What implications for magnitude judgments in bilinguals at different stages of language acquisition? *Cognitive Processing*, *17*(3), 225–241. https://doi.org/10.1007/s10339-016-0762-9

Vander Beken, H., & Brysbaert, M. (2018). Studying texts in a second language: The importance of test type. *Bilingualism: Language and Cognition*, *21*(5), 1062–1074. https://doi.org/10.1017/S1366728917000189

Venkatraman, V., Siong, S. C., Chee, M. W. L., & Ansari, D. (2006). Effect of Language Switching on Arithmetic: A Bilingual fMRI Study. *Journal of Cognitive Neuroscience*, *18*(1), 64–74. https://doi.org/10.1162/089892906775250030

Volmer, E., Grabner, R. H., & Saalbach, H. (2018). Language switching costs in bilingual mathematics learning: Transfer effects and individual differences. *Zeitschrift Für Erziehungswissenschaft*, *21*(1), 71–96. https://doi.org/10.1007/s11618-017-0795-6

Wang, Y., Lin, L., Kuhl, P., & Hirsch, J. (2007). Mathematical and Linguistic Processing Differs Between Native and Second Languages: An fMRI Study. *Brain Imaging and Behavior*, *1*(3–4), 68–82. https://doi.org/10.1007/s11682-007-9007-y

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, *8*(3), 231–256. https://doi.org/10.1162/jocn.1996.8.3.231

Wicha, N. Y., Dickson, D. S., & Martinez-Lincoln, A. (2018). Arithmetic in the Bilingual Brain. In *Language and Culture in Mathematical Cognition* (pp. 145–172). Elsevier. https://doi.org/10.1016/B978-0-12-812574-8.00007-9

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251. https://doi.org/10.1016/0010-0285(92)90008-P

Xenidou-Dervou, I., Gilmore, C., van der Schoot, M., & van Lieshout, E. C. D. M. (2015). The developmental onset of symbolic approximation: Beyond nonsymbolic representations, the language of numbers matters. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00487

Xenidou-Dervou, I., van Atteveldt, N., Surducan, I. M., Reynvoet, B., Rossi, S., & Gilmore, C. (2023). Multiple number-naming associations: How the inversion property affects adults' two-digit number processing. *Quarterly Journal of Experimental Psychology*, 17470218231181367. https://doi.org/10.1177/17470218231181367

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Ziegler, J. C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical Words are Read Differently in Different Languages. *Psychological Science*, *12*(5), 379–384. https://doi.org/10.1111/1467-9280.00370

Zuber, J., Pixner, S., Moeller, K., & Nuerk, H.-C. (2009). On the language specificity of basic number processing: Transcoding in a language with inversion and its relation to working memory capacity. *Journal of Experimental Child Psychology*, *102*(1), 60–77. https://doi.org/10.1016/j.jecp.2008.04.003

General Discussion

8 GENERAL DISCUSSION

8.1 Studies summary

Study 1 is a developmental study investigating transcoding in German French bilinguals of different age groups reflecting increased proficiency in French. German is the first language of mathematical acquisition (LM1) and French the second (LM2), see Table 1 of the general introduction. The task was either an auditory number-matching task or a number-naming. In the first task, a number is heard followed by four possible Arabic numerals (i.e. /Quatre-vingt-sept/ → 97 or **87** or 78 or 86), correct response in bold). In the second task, numerals have to be named (i.e. 87 → “Quatre-vingt-sept”). Both tasks were done in German and French, using a selection of numerals from 20 to 99. All groups recognized and named numerals French LM2 base-10 numerals (i.e. ‘30s to ‘90s) faster than base-20 numerals (i.e. ‘70s to ‘90s). All groups were also slower in the LM2 than the LM1 (i.e. LM2 cost) in both number naming and verbal-visual matching tasks after z-score standardization. Hence suggesting French vigesimal numerals (70’s, 80’s and 90’s) opacity are harder to transcode than more transparent base-10 numerals. The results also suggest the LM2 cost does not seem to recover with increasing proficiency in LM2. In sum, there is a lexical cost for the LM2 and for more opaque number word structures. Since increasing LM2 proficiency does not affect the effect of transparency and LM2 cost (i.e. we have similar results for younger age groups than adults), the following studies focused on adult bilingual samples, assuming the results would hold back to younger Luxembourgish populations, see Table 3.

In **Study 2**, we further investigated the question of LM2 cost in adults by replicating part of Study 1 (auditory-visual matching) and adding German and French monolingual groups as a comparison. To investigate the effect of morpho-syntax we used a similar auditory-visual matching design than in study 1. In Study 2 however here we manipulated the order of appearance of the unit or the ten parts of the two-digit numerals to visually match either the French ten-unit morpho-syntactic position (i.e. 4_ → 42) or the German inverted unit-ten morpho syntactic position (i.e. _2 → 42). For example, when a participant heard “forty-two”, this was either followed by a simultaneous (42), or two sequential conditions: unit first (_2) and ten first (4_) among four distractors. The results replicated the LM2 cost of study 1. In comparison to German monolinguals, we found generally slower responses for bilinguals in German (i.e. LM1). These results suggest a bilingual lexical cost for the LM1, that is bilinguals are generally slower than monolinguals. Finally, comparing a model based on French monolinguals with bilinguals in French (LM2) revealed an intriguing patter of results. While the unit-first condition mimicking LM1 number word position has relatively interfered with responses in LM2, the ten-first condition mimicking LM2's structure was relatively facilitated. In sum, in study 2 we find that bilingual's morpho-syntactic influence on lexical access in LM2 is malleably affected by LM1 and LM2's morpho-syntax, see Table 3.

In **Study 3**, we wanted to investigate if the lexical cost observed in Studies 1 and 2 also affects semantic number processing. We used a Priming Distance Effect Design (PDE) design (see Reynvoet et al., 2002). Participants had to name an Arabic numeral, the target. The target was preceded by a short (51 ms) number word masked prime in German or French. The PDE predicts that distance between the prime and target affects the naming time of the target (i.e. distance 3: Prime = "two", Target = 5, is named slower than distance 1: Prime = "four", Target = 5). Differently from Studies 1 and 2, we used exclusively single-digit numerals here (i.e. 1 to 9). Primes were either (randomly presented) number words in German or French, and Targets were to be named in German or French (blocked). The results show a similar repetition priming effect with German and French number words (i.e. Prime = "five"/"cinq", Target = 5 hence distance = 0). Hence suggesting that participants accessed LM2 number words' primes equally in both languages. The PDE, used here as an indicator of number's semantic access, was however found only with German number words primes and not with French number words. Hence suggesting the LM2 number words in French are less semantically associated than LM1 German number words. In sum, study 3 suggests that lexico-semantic associations of numbers are stronger in the LM1 than the LM2, see Table 3.

Table 3

Summary of the three studies with the results

Study	Paradigm	Design	Research Question	Group/Sample(s)	Stimuli	Results
1	Number naming & Number matching	Within ID Between Lang.	How does increasing LM2 proficiency affect morpho-syntactic mediation of lexical access?	Children & adults Bilingual 26 (5th g); 28 (8th g); 25 (11th g); 20 (adults)	2-digit Auditory ↔ Visual (AN)	LM2 cost Slower base-20 vs base-10 in LM2
2	Auditory-visual number matching	Within & between ID and Lang.	How does the morpho-syntactic decade-unit position influence German and French lexical access in monolinguals vs. bilinguals?	Adult Bilingual & Monolingual 55 (MonoDE); 56 (MonoFR); 50 (Bilinguals)	2-digit Auditory → Visual (AN)	LM2 cost Bilingual lexical cost Malleable morpho-syntax processing in LM2
3	Priming Distance effect (PDE)	Within ID Between Lang.	How does lexical and lexico-semantic access compare in bilinguals?	Adult Bilinguals 32 (Bilinguals)	1-digit Visual (AD) → verbal (NW)	LM2 cost Repetition priming in both languages PDE only with LM1 primes

Notes. ID = Participants, Lang. = Languages. g. = grades. MonoDE = Monolingual German. MonoFR = Monolingual French. AN = Arabic Numerals. NW = Number Words. LM1/2 = first/second language of math acquisition, see Table 1 of the general introduction. Levels of language influence on bilingual: Lexical access. Morpho-syntactic influence on lexical access. Lexico-semantic access.

In a nutshell, the three studies have shown lexical, morpho-syntactic and lexico-semantic influences on bilingual verbal number representations. In the following three chapters, I will detail different theoretical accounts dissected onto these three levels. In the first chapter, I discuss the *how* of the results, by which I mean the mechanistic cognitive accounts for number processing in bilinguals (§ 9 Cognitive theoretical interpretations and implications). The discussion of *why* the cost is found in LM2 but not LM1 is reserved for the two following chapters. In the second chapter, the effect of these three levels on bilingual number representations is discussed in terms of cognitive models for bilingual number processing on the associations of bilinguals' two verbal, visual and semantic representations of numbers (§ 10 Bilingual multiple number representations). Since one undiscussed assumption of those models is to distinguish balanced from unbalanced bilinguals, in the third chapter I attempt an integrative description of several factors of bilingual number processing that could affect long-term memory of numbers to account for proficiency (§ 11 Bilingual effects on number processing).

9 Cognitive theoretical interpretations and implications

In this chapter I will start by discussing the LM2 **lexical cost**, or why in all three studies (*study 1, 2 and 3*) performances were slower in French (LM2) compared to German (LM1) (§ 9.1 LM2 Lexical Cos). Then the **morpho-syntactic** effects, such as the effect of the base-20 transparency of power for the '70s to '90s France's French number words for transcoding in French' LM2 in *study 1* (§ 9.2 Language transparencies' morpho-syntactic modulation). In the same sub-chapter, I will discuss the effect of morpho-syntactic transparency of order of German number words (i.e. ten-unit inversion) on processing in French LM2 (*study 2*). Finally, in the third sub-chapter, we will discuss some cognitive accounts to explain the weaker LM2 **lexico-semantic**, as found in *study 3* (§ 9.3 LM2 lexico-semantic cost).

9.1 LM2 Lexical Cost

Slower responses in French (LM2) than in German (LM1) are found robustly in all three studies (see Table 3), suggesting a cost for lexical retrieval in the LM2 compared to the LM1. Since the LM2 cost is found when passing both from visual to verbal (number naming) and verbal to visual (auditory-visual matching task), it suggests that its origin comes from a general mechanism that is not task specific. Table 4 shows the LM2 cost for Luxembourgish bilinguals (see Table 1) in terms of ms, across different studies and tasks.

Table 4:
LM2 cost across studies

Study	Task	LM1	LM2	LM2 cost (LM2-LM1)
3	Number naming	642	665	23
2	Verbal-visual matching	1161	1266	105
1	Verbal-visual matching	795	879	84
1	Number-naming	697	855	158
Van Rinsveld et al., 2015	Simple additions	915	1060	145
	Complex additions	2740	3420	680

Notes: Data for Study 1: § 7.2.1 S1. Reaction Times (in ms) and § 7.6.2 S4. Verbal-visual matching, only adults and 30's to 50's numerals. Study 2, see Table 4, only simultaneous condition considered here. Study 3, see § 4.2 Filler prime, only the single digit no prime condition is considered here. Data from Van Rinsveld et al. 2015 manually extracted (hence approximated) from Figure 2 only adults with (*https://apps.automeris.io/wpd/*) and note that visual and auditory tasks of the studies are aggregated.

Hence, we can make several conclusions about the observed LM2 cost. First, the positive correlation between number transcoding and arithmetic of study 2 (see § 4.5.4 Correlation with arithmetic for bilinguals), suggests that the LM2 cost of study 1 to 3 might involve similar fundamental mechanisms as for solving arithmetic (see Van Rinsveld et al., 2015). Second, the LM2 cost **increases with task difficulty:** from single-digit naming (i.e. in Study 3), two-digit numbers (i.e. Studies 2 and 1) to complex additions (i.e. Van Rinsveld et al., 2015), see Table 4. Note that this increase seems to be rather exponential than linear (i.e. from 23, 158 and 680 ms). Hence small effects observed with simpler tasks could generalize and be amplified to complex ecological tasks such as doing arithmetic. Third, study 1's result on standardized scores[17] suggests the LM2 cost remains stable across age groups, hence after controlling for age groups' difference in variance, the LM2 cost does not resorb with increasing LM2 proficiency. In sum, the LM2 cost is robustly found across different tasks (auditory-visual matching and number naming), stimuli (single- and two-digit numerals), and age groups/levels of LM2 proficiency after controlling for variance (children and adults), see Table 4.

Several theoretical accounts that could explain an LM2 cost. I will start with two accounts that are important to consider for bilingual investigations but that can be excluded here. These accounts stem from the stimuli themselves and lower processes. The LM2 cost originates from **language surface differences** (which is a confounder in bilingual research, see Ellis & Hennelly, 1980). However, German number words are on average longer than French ones on average (see stimulus list of the three studies in the supplementary material). Also, Study 1 replicated the lexical when only four-syllable long number words were taken into consideration in the analyses. The LM2 cost could also stem from **low-level processing stages** such as comprehension or production of French numerals. For example, it could be that more time is required for reading (in Study 3), phonological decoding (in Studies 1 and 2) and articulating (in Studies 1 and 3) numerals in French than in German. For reading it could be that bilinguals were less familiar with French grammar, word form or bigram frequency. For phonological decoding, it could be that low-level speech processing is highly specialized in German. On the production part, it could also stem from motor planning and coordination of the articulation of words, which would be slower in French than in German. Although there is no direct data to exclude the phonological decoding and articulation account in studies 1 and 2, one experimental result of study 3 excludes the possibility of slower reading access to number words in French. In study 3, we have found a repetition priming of French number words on Arabic digits (i.e. the prime “cinq” facilitates the naming of 5 into /cinq/ or /fünf/). Hence, even a very short presentation (priming) of a number word in French, followed by a target of the same number could facilitate its processing. This suggests that not only LM2 number words are decoded (i.e. reading) but that they were activated up to the lexical level. Furthermore, the result that French number words can also facilitate the naming of Arabic numerals in German (i.e. the

[17] See in study 1’s: Fig 3. Z-score reaction times of the reading aloud task**.** and § 7.4.1.1 S2. RT z-score. See also: Fig 6. Z-score reaction times of the verbal-visual matching task and § 7.4.2.1 S2. RT z-score.

prime “cinq” facilitates the naming of 5 into /fünf/) suggests the lexical access from French number words is general and co-activates the equal lexicon in the LM1. In sum, since LM2 shortly primed number-words are lexically activated, the LM2 cost must stem from a difference of association with the lexical verbal association between LM1 and LM2.

Now we will see some cognitive accounts that could explain the observed LM2 cost as a slower lexical retrieval in LM2 than LM2. One of them is the **language competition** which suits particularly well to explain Arabic numeral naming performances. Since Arabic numerals are the same symbols for both languages, they could co-activate both LM1 and LM2 lexical representations (i.e. 5 → “fünf” & “cinq”). Because both languages are co-activated but only one can be named, one of the languages needs to be inhibited. According to the IC model (Green, 1998), the language or words with weaker associations (i.e. LM2) are inhibited more efficiently, leading to facilitation for accessing the language or words with stronger associations (i.e. LM1), see § 4.2.3 Inhibitory and Adaptative Control (IC and AC). Similar prediction is made by connectionist models such as the BIA+ and RHM model (see respectively § 4.2.1 Revised Hierarchical Model (RHM) and § 4.2.2 BIA+, BIA-d and Multilink). For the RHM model, this is due to weaker lexical associations of the L(M)2 compared to the L(M)1. For the BIA+ model lower frequency of use of number words in L(M)2 than in L(M)1 leads to slower or less efficient processing of the lexical form. In sum, weaker LM2 lexical associations than LM1 are at the origin of the LM2 cost.

Note that the language competition account can also explain the **bilingual lexical cost**: the results of study 2 where bilingual LM1 (German) was slower than German monolinguals. When bilinguals hear the German number word “Zwei-und-Vierzig” it might co-activate the French “Quarante-deux” which needs an additional inhibitory system compared to monolinguals leading to a slightly slower process. Note that, alternatively it might be the Luxembourgish number word form at the origin of the language competition. For example, it might be that the bilingual lexical cost originates from language competition between German and Luxembourgish, rather than German and French. Luxembourgish coactivation is plausible for being closer to German (i.e. phonologically, and orthographically). Alternatively, the bilingual language cost could also be due to the absolute frequency of L(M)1 since being exposed to two languages can lead to on average less exposure than monolinguals (see Mägiste, 1979).

9.2 Language transparencies' morpho-syntactic modulation

Study 1 and **Study 2** results show different influences of language morpho-syntactic properties on number's lexical access in bilinguals. **Study 1** mainly investigated the transparency of power of the French mixed number system, where while number word s below 60 follow a base-10 system, number words between 70's and 90's are in base-20. S**tudy 2** results regard transparency of order in German, such two-digit numbers above 20 in German are inverted compared to the ten-unit place value system of Arabic numerals.

Study 1 shows that in addition to the LM2 cost, a cost for the base-20 number words in French (i.e. '70s to '90s), an effect of **transparency of power**. The result of study 1 replicates previous results, which however have mostly investigated in monolingual (Camos, 2008; Saad, 2010; Seron & Fayol, 1994) and bilingual children (Van Rinsveld, Schiltz, Landerl, et al., 2016). Study 1 results are however found in bilingual adults. French morpho-syntactic for those numerals have a different base (i.e. base-20) that does not match Arabic numerals (i.e. base-10), nor French number words between '30s to '60s (i.e. base-10). In addition, more than half of the number words between the '70s and '90s also include irregular teen numerals (i.e. 71 to 76, literally "soixante et *onze*" to "soixante-*seize"* and 91 to 96, literally "quatre-vingt-*onze*" to "quatre-vingt-*seize*"). All those factors combined increase the opacity of France's French '70s to '90s number words. As in the case of lexical cost described above, we can exclude that slower responses for base-20 originate from differences in number word length compared to base-10 number words. Study 1 resulted in slower reaction times even after controlling for number word length (see § 7.4.1.2 S2. Subset data: four-syllable length and § 7.4.2.2 S2. Subset data: four-syllable length), suggesting it is independent of number word length. The underlying mechanisms that make base-20 numerals in French more difficult to transcode than base-10 numerals might be similar to inverted numerals since they involve morpho-syntactic properties of these languages.

Study 2 investigated the question of **transparency of order** with an experimental approach to answer the question of how units and ten morpho-syntactic positional order are processed in LM1 and LM2 compared to monolinguals. The results show that LM1's morpho-syntax modulates number processing in a similar way to bilinguals in German than German monolinguals, given non-significant differences between monolingual's ten-unit order. In contrast, for bilinguals in French (LM2) the order of ten or unit presentation modulated the responses. Priming the unit part, mimicking LM1's morphos-syntax, interfered with LM2 processing (i.e. slower responses). On the other hand, priming the ten-part, mimicking LM2's morpho-syntax, facilitated processing in the LM2. These results suggest a cross-language transfer from the LM1 inverted morpho-syntactic processing over the LM2 but also an LM2 morpho-syntactic transparency advantage over the LM2. Such that the LM1 morpho-syntax interferes with LM2 (i.e. the unit-fist condition interferes with processing in French) but also the LM1 transparency facilitates processing LM1's morpho-syntax (i.e. the unit-fist condition facilitates processing in French). All this however relative to the bilingual lexical cost and LM2 cost. In sum, the LM2 might be more malleable regarding morpho-syntax processing than the LM1.

Common cognitive mechanisms are probably underlying the morpho-syntactic modulations of the LM2 found in Studies 1 and 2. The mechanism's interpretations differ depending on postulates from different **models of numerical representations** (see § 2 Models of numerical representations). For the abstract model of McCloskey (McCloskey, 1992; McCloskey et al., 1985), the effect of transparencies could affect the comprehension and production stage. For example, since the model assumes number-word comprehension requires the decomposition in terms, since the input is less transparent this decomposition might be computationally heavier than for more transparent number words. Similarly for production, where several rules are required. For Power and Dal Martello, (1990)'s model the morpho-syntax would affect the semantic representation of numbers. Hence the verbal number words morpho-syntactic structure would affect how numbers are semantically represented.

For asemantic models such as the Triple Code Model (**TCM**) of (Dehaene, 1992) morpho-syntax would affect the processing of verbal representations of numbers as it would for other words and does not impact semantics. Indeed a procedural account for number transcoding would be enough to explain the effect of morpho-syntax on number transcoding tasks. Indeed, transcoding opaque numbers requires a greater number of rules for transcoding which could slow down the association between visual and verbal code to process and develop. This account would be sustained by Dotan and Friedmann, (2018) model and ADAPT model (Barrouillet et al., 2004) that I will describe in more detail.

Dotan and Friedmann's (2018) model for number reading (i.e. 42 → "Zwei und Vierzig") distinguishes the visual analysis processes from the verbal production process. Although initially, the model suggested morpho-syntactic language differences would impact the verbal production stage such as in lexical retrieval and in forming the word frames. Later research has led to the possibility of top-down language influence on the serial order scanning of the visual analyser (Dotan, 2023). For example, German speaker might start scanning two-digit numbers from right-to-left. Indeed accounts of the compatibility effect report that reading direction impacts ten-unit inversion (Moeller, Shaki, et al., 2015). However, if we did not observe an effect of the unit first mimicking the inversion in German for monolingual German and bilinguals in German of study 2. This could either be due to a methodological issue (i.e. the primes were too long to elicit such an effect) or it might be explained in a developmental term such as that those effects are observable in children (or neuropsychological cases) but are not large enough to elicit a difference in adults performances.

The **ADAPT model** suggests that children transcode numerals by applying procedural rules which are specific to language morpho-syntax. Concrete examples of these rules are described with precision in the model, which makes it also computational. Since more opaque languages require more rules, they lead to more errors and slower responses for children. The developmental part of ADAPT predicts that adults retrieve complex number words such as base-20 and invert directly from long-term memory as lexical units. Therefore bypassing procedural rules. This automatization, which reminds the "lexicalization" in Deloche and Seron (1982), would be led by increasing frequency of retrieval. For example "quatre-vingt-dix-neuf" or "zwei-und-vierzig" would be single lexical items corresponding to 99 and 42 which are directly retrieved from long-term memory, as words outside the numeral's lexicon. However, this prediction is not met for the LM2. In Study 1, we find an effect of transparency of power in adults. In Study 2, monolingual German and bilingual in German (LM1) are not impacted by the unit how the unit-first condition than French monolinguals. Bilinguals in French (LM2) were relatively affected by the Arabic numeral's experimental ten-unit structure mimicking either the properties in German (unit first) or French (ten first). While the ADAPT proposition that number words are directly retrieved from long-term memory is met for LM1, it does not fit our results on LM2.

9.3 LM2 lexico-semantic cost

Our results of a weaker lexico-semantic connection in the L(M)2 **contrast with other findings**. For example, the Priming Distance Effect (PDE) was found in L(M)2 (Duyck & Brysbaert, 2002), with Prime Arabic numerals and Target number words to be named and translated. However, it might be that the priming effect emerges there given the longer naming times in the L2, leading to a larger SOA (Van den Bussche et al., 2009). Another explanation could be that the PDE we observed is mediated by number word reading, such that L(M)2 number words are slower to read than L(M)1 (Coltheart et al., 1993). Therefore, rather than being a representational difference, it would be a difference in reading speed. One of our results in Study 3 however refutes this argument. We found a repetition priming when we presented an LM2 number word prime followed by an Arabic numeral target named in both languages (i.e. Prime = “trois”, target = “3” to be named /trois/ or /drei/). In other words, we observed a facilitation in the LM2 number word on the Arabic numeral representing the same. Hence this facilitation can only be explained in that the prime was read and there was a lexical access. Hence for the semantic mechanisms for the absence of a PDE in LM2 found in Study 3, I will discuss two theories for the origin of number’s semantics: the DSS and ANS.

The Discrete Semantic System (**DSS**, Krajcsi et al., 2016; Sella et al., 2021) predicts that semantic effects such as the PDE would arise from the discrete association between symbolic numerals. Hence the number distance effect would arise from the stronger association across closer than more distant number words (i.e. “four” <-strong-> “five” and “two” <-weak-> “five”). A stronger association for close than large numbers would arise because they are more likely to be retrieved together than more distant numbers. For example, each time we count (i.e. ...“two”, “three”, “four”, ”five”) the co-activation of closer numbers (i.e. “four”, ”five”) would consolidate their association. Note that this consolidation would therefore follow a Hebbian learning principle.

For the second account the semantic associations that explain the PDE are that numerals are mapped onto an abstract asymbolic system, the Approximate Number System (**ANS**). The ANS indeed predicts smaller overlaps of small number representations than larger ones (i.e. “two” and “five would have a smaller overlap than “four” and “five). Considering this account, the LM2 weaker lexico-semantic associations would arise from a poorer mapping of LM2 number words with the ANS compared to the ANS. Since the ANS account relies on the Triple Code Model (TCM) which stipulates different neuro-cognitive modules for each module, this could be directly tested. For example, future studies using neuroimaging techniques could test if co-activation of the left temporal (where the verbal code should be principally processed) and bilateral intraparietal sulcus (where the ANS should be processed) is stronger with LM1 than

LM2. In sum, both the DSS and ANS accounts make almost undiscernible predictions on the results for tasks such as the PDE.

10 Bilingual multiple number representations

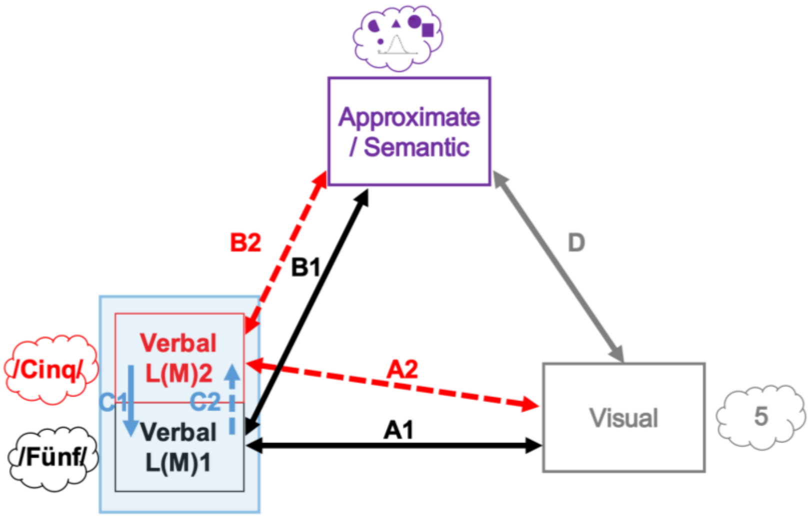
We have discussed several cognitive mechanisms that can explain the different effects observed in bilingual number transcoding assuming a language dominance for the L(M)1. I will now attempt to discuss a model to account for the interplay between L(M)1 and L(M)2 and then several theoretical accounts that could explain the origin of the differences between LM1 and LM2.

10.1 Bilingual Triple Code Model

The Bilingual Triple Code Model (**BTCM**) has been developed along with Study 3 (Lachelin et al., 2023). A previous proposition analogous proposition of a bilingual triple can model can be found in the discussion of the PhD thesis Van Rinsveld, (2015). These propositions are inspired by the Triple Code Model (TCM, see § Triple Code Model (TCM 2.2) by adding a second verbal code. In the BTCM there is an approximate semantic code which allows for non-symbolic estimations without languages such as in pre-verbal children. Symbolic codes have independent lexico-visual (i.e. A1 and A2, see Figure 8) and lexico-semantic (i.e. B1 and B2, see Figure 8) associations with each language. The verbal code associations are theoretically asymmetrically, such that the association is stronger from L2 to L1 than from L1 to L2 (i.e. C1 and C2). The models' independence of lexico-lexical (i.e. C1 and C2), lexical (A1 and A2) and lexico-semantic (B1 and B2) associations indicates that these associations can differ in strength. The difference in the strength of the associations for the verbal codes depends on individual bilingual language profiles.

Figure 8

Bilingual Triple Code Model



Notes: L(M)1 = first Language (of Mathematical) learning, L(M)2 = second Language (of Mathematical) learning. Blue unidirectional arrows indicate translations between the two languages existing for the verbal code: forward translation from L(M)1 to L(M)2 (C1); backward translation from L(M)2 to L(M)1 (C2). Dashed arrows indicate weaker associations compared to full arrows. The arrows correspond to: bidirectional lexico-visual associations with L(M)1 (A1), bidirectional lexico-visual associations with L(M)2 (A2), semantic access from - and to - L(M)1 (B1), semantic access from - and to - L(M)2 (B2), independent semantic access from - and to - the visual code (D).

The BTCM therefore postulates the independence of the lexical and lexico-semantic associations between the language codes of both bilingual languages. Note that the independence of lexical and lexico-semantic associations concerns the independence of access to these representations. However, underlying syntax and procedural rules can transfer across languages (such as the cardinality principle for example see Wagner et al., 2015).

While the BTCM postulates different verbal representations in bilinguals, these could also differ visually as with **different notations** or writing styles (as in the encoding complex model, see Campbell, 2005). For example, Japanese can be written in Kanji (i.e. ご logographic) or Kana (五 phonological), or Arabic (i.e. 4 = ٤, 5 = ٥). In that case, the language lexico-semantic independence for two verbal codes could also apply to distinct visual codes. This in turn would depend on the overlap between bilingualism and bilateralism, the acquisition and proficiency in different character systems.

10.2 Bilingual lexical cost in L(M)1 and L(M)2

The difference leading to slower lexical retrieval in L(M)1 compared to L(M)2 is accounted for in the BTCM. The origin of the lexical cost could stem from frequency of use, Age of Acquisition.

The difference might be due bigger frequency of use and exposure to the LM1 than LM2 (see for example the weaker link hypothesis in bilinguals Gollan et al., 2008). For example, a larger **frequency of retrieval** of number words in German than in French might have built stronger lexical associations between Arabic numerals and LM1 than LM2 number words. If we consider the Luxembourgish samples investigated in the three studies, it could be that despite mathematics are thought in both languages, German and Luxembourgish are preferentially used for numerical tasks such as mental calculations therefore increasing the frequency of retrieval and associations of LM1 compared to the LM2 (see also § 11 Bilingual effects on number processing proficiency). This account would fit the BIA+ account and could also explain the bilingual lexical cost. Remember that in BIA+ the activation level depends on exposure. Bilinguals can only be less exposed to the LM1 than bilinguals on average since contrary to monolinguals their exposure time is divided across two languages.

Besides the larger frequency of lexical retrieval for German number words compared to French ones described above, the difference might come from other sources such as later **Age or Order of Acquisition** (i.e. that French number words are formally acquired and consolidated later than German ones). For example, since the LM1 is acquired earlier than the LM2, it might have led to privileged access due to a critical period of language acquisition (see for example the critical period hypothesis, CPH in Weber-Fox & Neville, 1996). Developmentally, since both LM1 and LM2 are acquired during an ageing period of brain maturation (Gogtay et al., 2004), it might be the brain maturation stage at which the LM1 is acquired is more auspicious for language acquisition than when the LM2 is acquired. As already noted earlier the frequency of use and critical period accounts are difficult to disentangle since earlier acquired languages also usually benefit from more frequency of retrieval.

10.3 Bilingual morpho-syntactic modulation of lexical access

Study 2 comparison of monolingual German and French in the auditory-visual matching task suggests both groups resulted in similar patterns of results for the experimental manipulation of transparency of order. This is shown by the non-significant difference when comparing the patterns of results from the three conditions of Study 2. Moreover, the comparison of German monolinguals with bilinguals in German (LM1) did not show morpho-syntactic modulation in lexical access. Hence, on the premises that Study 2's method is sensitive enough, the unit-first did not affect these three cases. For example that number-words are directly retrieved from long-term memory as predicted for example by the ADAPT model (Barrouillet et al., 2004). For the LM2 however, Studies 1 and 2 suggest some evidence for a morpho-syntactic modulation. Suggesting morpho-syntax affects lexical access in adult LM2.

A way to reconcile our results with the LM2 and **ADAPT** model could rely on **practice**. Adult's LM2 relative frequency of retrieval would not attain the same level as monolinguals to bypass procedural processing and become automatized. Hence LM2 practice might be important for equal proficiency in both languages (see for example Hartshorne et al., (2018)). For example, it might require around 10.000 hours of practice and hence repetitions regarding acquiring performance expertise (Ericsson et al., 1993). Before that threshold, number words would therefore be processed procedurally.

A **procedural account** could speculatively explain why in Study 2, the unit-first condition mimicking German LM1's influences number processing in the LM2. This would result from procedural training of using the unit first in German, hence affecting a general number-word recognition mechanism (see for example ADAPT's transcoding rules). In other words, when the Luxembourgish bilinguals hear a number word in French, they automatically activate the procedural process which is useful for inverted number words, hence focusing on the unit. On the other side, the French ten-first morpho-syntax seems also to be automatized to some extent, such that the ten-first condition facilitates processing in French's LM2. However, interestingly the automatization of the French ten-first procedure does not seem to affect transcoding in LM1. If this was the case we should have found that bilinguals in German interfered in the first conditions, which we did not find in our results.

This account would fit with procedural theories of language learning. The **procedural declarative model** (DP Ullman, 2004) suggests that the mental lexicon is supported by declarative memory in the temporal lobes. While procedural memory sustains motor and cognitive skills for sequences in a specific network (composed of frontal, basal ganglia, parietal and cerebellum). Hence LM1 would be fully integrated into the declarative model, while LM2 number words would still require procedural mechanisms. For language acquisition in general and second language acquisition in bilingualism, it has been suggested that similar mechanisms as nonlinguistic motor learning and language learning (see Hernandez, 2013). Both require long training for expertise and have specific AoA. AoA has been proven to play a role in procedural perceptual learning. For example in perceptual narrowing, the fact that with developments babies restrict their phoneme perceptive range to the phones they hear. In Japanese for example the phoneme /r/ is inexistent and hence adults cannot discriminate it from /l/ (i.e. Miyawaki et al., 1975).

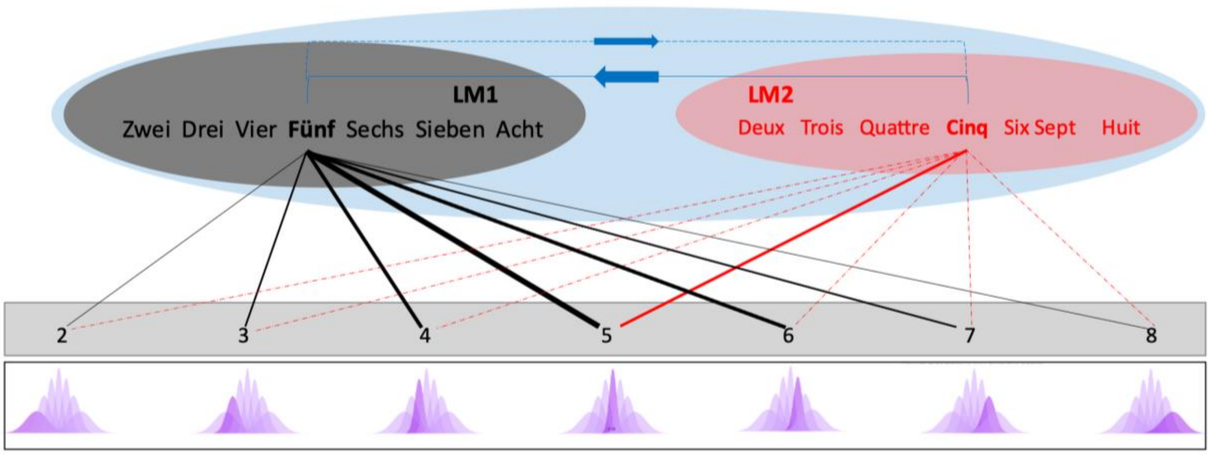
10.4 Bilingual lexico-semantic associations

The LM2 lexico-semantic effect found in Study 3 on the other is accounted for by the BTCM, such that LM2 lexico-semantic associations are weaker in the LM2 than LM1. In the BTCM it would mean (1) the number word primes activate the verbal representations, and (2) the Arabic numeral targets activate the visual representation of numbers. At this step, both LM1 and LM2 commonly activate the same Arabic numeral (i.e. “fünf”/”cinq” → 5), but possibly slower in LM2 than LM1 number word primes. Then (3) the co-activation of the semantics occurs more efficiently with LM1 than with LM2 primes. Finally (4) the target Arabic numeral is read in the required language.

Complementary to a modular model such as the BTCM, it nevertheless also possible to conceive a connectivist model such as depicted in Figure 11. In this kind of model, each lexical element (i.e. verbal number words and visual Arabic numerals) could be represented as lexical discrete representations. The first form of association is lexical associations between languages such that “Fünf” is associated with “Cinq”. This association is asymmetrical, such that it is stronger from LM2 to LM1 than from LM1 to LM2 (see blue arrows in Figure 11). The second form of association is between verbal and visual codes, such as “Fünf” ↔“5” and “Cinq” ↔“5”. These associations might be built very fast and early in development. The third form of association requires however more time to build and develop, it is the lexico-semantic associations such that numbers are associated with neighbouring numbers. These lexico-semantic associations are relative to the distance with other numbers. A possibility of how lexico-semantic associations work is represented in the bottom panel of Figure 11. Each lexical form would co-activate neighbouring lexical items, for number the strength of this co-activation depends on the distance (and likely size too). For example “fünf” and 5 would lead to a maximal overlap of lexico-semantic activation, while “Fünf” and 2 to smaller overlap, hence a smaller and more distributed activation.

Figure 9

Connectivist account for bilingual number representations and semantics



Note: The blue area indicates an integrated vocabulary. The grey area contains the verbal number representations of LM1 (i.e. German) and the red areas of LM2 (i.e. French). Number words and Arabic numerals are exact symbolic representations. Note that Fünf and Cinq are associated, such that they can be directly translated without involving semantic levels (this might be asymmetric). The Violet bell curve represents representations of co-activations (i.e. fünf and 5 co-activation overlap is stronger with 4 and 6 than 2 and 8).

This model would be parsimonious in that these associations would only need the frequency of use and exposure of each language. Indeed lexical associations across languages occur more often between the same numbers. For example, language translation oft involves finding the lexical equivalent in the other language: “Fünf” ↔ “Cinq”. These translations might occur more oft from the LM2 to the LM1 since at least at the beginning of development they might be necessary to access semantics- The lexico-lexical associations between visual and verbal would occur each time a number is named (i.e. 5 → “Fünf”). Finally, lexico-semantic associations might occur during counting as well as language frequency, size effect might for example be only explained by number words language frequency (Dehaene & Mehler, 1992)..

11 Bilingual effects on number processing proficiency

In the previous two chapters, we have seen different effects of weaker associations of the LM1 compared to the LM2. In the following I will discuss four concrete main cognitive dimensions that have potentially impact the processing of numbers in bilingual individuals: on the language level (§ 11.1 Linguistic properties (“lingualism”)) on the individual level (§ 11.2 Bilingual language profiles) and on the context level (§ 11.3 Language learning and testing context). Finally we will see these level can lead to complex interactions (§ 11.4 Complex Interactions). Note that besides the cognitive accounts discussed here for differences between LM1 and LM2, there might be other theoretical accounts that might affect bilingualism such as ethnolinguistic vitality and prestige and intelligibility between both languages (Martini, 2021). For example, the LM1 might be associated with more societal power leading to more institutional support and a better status perception (i.e. prestige).

11.1 Linguistic properties (“lingualism”)

Differences in language properties of bilingual languages might differently affect second language acquisition. Property differences could be for instance syntactic, lexical, orthographic or phonological. The more those linguistic properties differ between languages, affect **linguistic distance**. Acquiring an L2 that is more distant than the L1 might be more difficult. Inversely close languages might be facilitated L2 acquisition. For example, Luxembourgish is linguistically closer to German than French since it shares more properties with German (see Martini, 2021). Hence part of the LM2 costs could be due to linguistic distance between German (LM1) and French (LM2). Hypothetically if the LM2 would have been linguistically closer to German such as Dutch the observed LM2 cost might be smaller, which would be explained by orthographic overlap between the languages (see for example Duyck & Brysbaert 2008). In addition to orthographic overlap, the morpho-syntactic structure of number words for two-digit numerals matches between German and Dutch (i.e. they are inverted). On the other way round, instead of an interference due to linguistic distance between German and French, there could be a facilitation by the linguistic closeness between Luxembourg and German.

Some of these linguistic properties might be more **hierarchically more important** in determining second language proficiency. Syntax is the structure of a language; the effect of syntax has mainly been investigated here with regards of morpho-syntactic aspects number transparencies More transparent languages facilitate the acquisition of number words. But also phonology or number word length might affect both learning (longer words are harder to learn) and retrieval (longer words are slower to retrieve) of a second language. Hence these characteristics are affected by the physical linguistic properties of languages: surface for grammar and phonology and structure for morpho-syntax. It would be interesting to know which language characteristics are more important to correspond between a known and learned language, such as if it is syntax, phonology or grammar. This question should be addressed by future studies in numerical cognition.

11.2 Bilingual language profiles: heterogeneities and homogeneities

Another important factor influence how numerical representations are accessed across languages is the language profile of each bilingual (see § 4.1 Bilingual heterogeneity). Bilingual language profiles are shaped for example by the timing of acquisition such as early and late bilinguals. In parallel, exposure affects the frequency of use of each language.

The second factor affecting memorization is use and exposure which in turn affects relative **language frequency.** More frequently a number word is retrieved the stronger its associations become. Associated with language frequency is **AoA**. Memory consolidation might be stronger when done in a certain developmental time window than the other, in the sense that the memory traces would be easier to consolidate at a younger age and would need more effort at an older age (cfr. CPH). For example early bilinguals might not differ in accessing verbal representations in both languages. Bylund et al., (2022) for example did not find a bilingual lexical cost in L1 for early bilinguals but for later learned languages, hence sequential bilinguals.

Bilinguals profiles are also determined by the language **context of acquisition and consolidation** the consolidation of specific number word might be context-dependent. For example, retrieving known numbers such as dates or a phone number might be easier in one language than the other depending on the learning and retrieval language context. Finally, the testing context also plays a role, depending on which language was pre-activated. This might act like language priming, when a certain language network is pre-activated it facilitates lexical retrieval in that particular language compared to the other. This pre-activation might be general (i.e. over a day or month) and specific (i.e. depending on the language used for the previous item).

Language profiles might also **interact with the language characteristics** of each language: learning a new language also depends on the language that is already known. Hence a new language with more linguistic similarities (closer or having evolved from a common language such as the romance languages) with the one already known are easier to learn than more distant languages. Finally, it is easier to learn a new language when another language has already been learned (Grosjean, 2010), which might also influence multilingual language profiles concerning language acquisition.

11.3 Language learning and testing context

The degree of language activation is also context-sensitive. By context I mean the language learning and retrieval context. For example, if the language of learning and retrieval is the same or switches. Or if a language switch occurs in the short term during retrieval or testing (i.e. switching the language of retrieval within a test). Or if there is a switch in the general language context (i.e. the language used before the test).

Language Switching Costs (LSC) are found when switching languages than when remaining in the same language both in the long-term between test and retrieval and in the short term within a test. LSC predict worse performances for switching than not switching languages. **Long-term LSC** resulting from the switch between the language of encoding and retrieval could be explained in that encoding is language-specific. This is particularly evident in training studies, such as for arithmetic (i.e. Spelke & Tsivkin, 2001b and Saalbach et al., 2013). A LSC cost could explain the difference in association weights found in the Luxembourgish sample investigated in the three studies. Indeed, the strongest associated language LM1 could be the language in which arithmetic and more basic number representations are learned. The weakest associated LM2 instead could be the language in which "higher" and mathematical content is acquired. Algebra, trigonometry, or differential equations rely on other principles than the "basic" principles of numbers such as distance and arithmetic. Hence, we can suppose that arithmetic is learned in German rather than in French leading to LSC when they are tested in French.

Language context can also change between items of a test leading to **short-term switching costs** (i.e. it would be easier to name 5 -> /fünf/ if preceded by the items 8 -> /acht/ than by 8 -> /huit/). Although this situation does not concern Studies 1 to 3, it is nevertheless a factor influencing the bilingual processing of numbers. Also, as suggested by the RHM, the effect of the language context on LSC might not be symmetrical between L(M)1 to L(M)2 *or vice-versa*. Hence suggesting that LSC interact with the language profiles. In other words, the effect of LSC depends on the status of the languages and the direction in which the switching occurs.

Finally, the **general language contexts** seem also to affect the pre-activation of a language such as suggested by the language mode theory. Hence a language that was previously activated (such as before the test) would be more facilitated compared to the other language. This can also affect arithmetic in bilinguals (see Van Rinsveld, Schiltz, Brunner, et al., 2016).

All those factors might predict the strength of association for lexical retrieval of number words in each language, however, their relation might not only be additive but also might lead to complex interactions.

11.4 Complex Interactions

By looking at a practical example it is easy to see how the previous three dimensions might also give rise to complex interactions among each other. For example to predict which languages would lead to better performances in an unbalanced proficient bilingual with an LM1 German who spoke only in French the previous week and is at the end of a test in French? Would the predictions differ if the home language or the LM1 was Portuguese (i.e. closer to French for example regarding transparency of power)? Indeed all the above-mentioned factors might interact.

These kinds of complex interactions might lead to hard-to-predict performances in uncontrolled but more ecological environments. They might also be the explanation for why so much research on bilingualism leads to contradicting results such as concerning the cognitive bilingual advantage or disadvantage. Moreover, it would need to be understood if these interactions are additive or multiplicative. If they are additive, it would mean bilingual performances could be predicted from monolingual data by adding different weights to account for the factors described above. However, if they are multiplicative then the predictions might differ for each interaction. The degree of influence of these processes furthermore depends on bilingual language profiles (i.e. proficiency, LM, etc.). All these interactions would need to be addressed in future research or by a unified model of bilingualism.

Nevertheless, I don't think bilingual performances are doomed under the unpredictability of complex interactions, nor that future researchers should be discouraged by this complexity. Some factors such as exposure and training are likely to be more relevant in predicting bilingual performances', for example, study 2 shows bilingual performances might be more unpredictable for L(M)2 than L(M)1. Also, group differences are nevertheless informative about bilingual number processing.

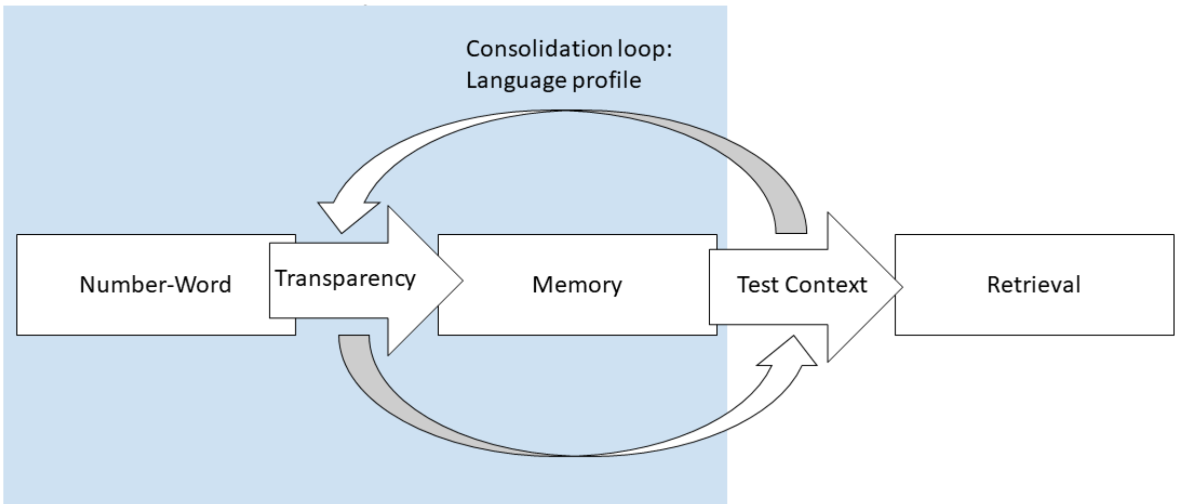
11.5 Bilingual long-term memory

I will attempt to break down the LM2 cost in terms of long-term memory (LTM) encoding consolidation and retrieval. Indeed the end LM2 cost result is a snapshot reflecting bilingual language experiences (i.e. exposure and use). In terms of LTM, this experience can be broken down into language-specific encoding, consolidation and retrieval.

If we consider the LM2 cost as an outcome of slower LTM retrieval, it could be explained by shallower encoding. For numbers, **encoding** could mean the mapping between different codes (i.e. verbal visual and semantic). This mapping is exact, meaning that each natural number has one – and only one - corresponding number word. If we consider the encoding of numbers semantic or arithmetic in the Luxembourgish school system, their encoding occurs principally in the primary years of education, hence in German. Therefore it could be that Luxembourgish bilinguals have encoded number representations in a content language-specific manner. Following the critical period hypothesis, it could also be that the quality of encoded number representations depends on age. **Consolidation** is another important component of LTM. As for encoding, consolidation might be language-content specific. For example, provided Luxembourgish students prefer to process numbers in German, they would tend to prefer German for processing numbers, hence consolidating associations of German number words. In general, bilinguals respond to the use of the most dominant language to solve arithmetic (Dewaele, 2007). Hence each time number words or arithmetics are retrieved in one language the individual relative frequency of this number word increases and consolidates the memory trace, see Figure 10. Concerning **retrieval**, the LM2 cost might occur for retrieving information from LTM. Besides being the results of shallow encoding and consolidation, retrieval processes themselves might be slower for the LM2 (such as the language competition account discussed above). In monolingual adults, the influence of linguistic characteristics on LTM lexical retrieval might diminish as the access to the mental representation of numbers becomes more automatized (Logan, 1988), meaning they are directly retrieved from long-term memory. This might however not be the case for bilingual's LM2, as represented by the effect of language transparency in Figure 10. Also, considering context-specific linguistic effect such as LSC and language mode would act at retrieval.

Figure 10

Model for bilinguals' long-term memory lexical retrieval of number words



Notes: Predicting bilingual lexical retrieval from a long term memory perspective. The blue shadow indicates the part that changes with development.

11.6 Limitations and constraints of generalization

All the bilingual samples investigated in these studies are characterized by having followed the bilingual Luxembourgish school curriculum. Hence to understand the degree those results might generalize to bilingualism in general we need to understand both the school system and socio-cultural backgrounds. Luxembourg is a multilingual country, officially Luxembourgish is the national language, and the legislation is in French, but the two languages as well as German are legal languages for judicial and administrative matters (*Loi Du 24 Février 1984 Sur Le Régime Des Langues. - Legilux*, n.d.). Luxembourgish German and French can be encountered in everyday life in Luxembourg: media can be found in all three languages: television and radio in Luxembourgish (i.e. *RTL - Radio*, n.d.) and journals in both French (i.e. *L'essentiel: Actualité Du Luxembourg et News Internationales*, n.d.) and German (i.e. *Luxemburger Wort | Luxemburger Wort*, n.d.). The actual amount of exposure someone living in Luxembourg had of these three languages is very difficult to assess and might vary from individual to individual. Besides the three languages described before, a large diversity in spoken languages can be found in Luxembourg, such that only 42% of the students speak Luxembourgish at home (*The Luxembourgish Education System – An Overview*, n.d.). This is due to Luxembourg's having one of the highest rates of foreign residents in the world, which is estimated at 47 % (*World Bank Open Data*, n.d.). Hence the socio-cultural profiles with which the students enter the Luxembourgish school system can be very diverse.

Hence *multi*lingualism would be a more appropriate term to describe this sample, however, I use the term *bi*linguals. This is because we investigated and compared only the two languages that are formally acquired for mathematics. Also, finding pure bilinguals is as difficult as finding pure monolinguals (for example having zero knowledge of English). The reason is that most investigated populations are rather multilinguals (i.e. use and comprehension of more than two languages), and most of the experimental evidence and comprehensive models are about two languages. In that sense, the term bilingualism here is used as a sub-seed and partial equivalent of multilingualism. It is pragmatically used for a comprehensive description of the underlying theoretical cognitive mechanisms and experimental evidence, such that more languages could be added to describe and understand the full complexity of the multilingual experience.

Now, with regards to the studies presented above this diversity might have some influences on the bilingual cognitive processes we have tried to assess. First since socioeconomic level affects students' education achievements. Secondly because of the linguistic background diversity of our samples. Indeed, most of the participants reported speaking multiple different languages, mostly adding English and Portuguese to Luxembourgish, German and French. We have tried to carefully control for the language backgrounds in the study which had the largest sample: study 2. Nevertheless, even in this study, we found a cost for French compared to German. Furthermore, all 2 studies use repeated measures or within-subject designs, meaning that all participants did the same tasks in German and French. The advantage of this within-subject design is that we can compare individual performances in both languages. Hence limiting the influences from socio-economic status and language diversity. With regards to the Luxembourgish school system, the studies capture the performances in a definite developmental time: young adult university students. This brings a double strength to our conclusion of an LM2 cost: first, these adults might benefit from a "recency effect": they have been exposed to using French to do mathematics in school in the past 7 years. Moreover, we also benefit from a "selection bias", in the sense that to access higher education those students must have already been good in languages and mathematics (in particular doing mathematics in French). In other words, the LM2 cost observed in these three studies should only be bigger in the general population following the Luxembourgish school curriculum. For example, students who do not have Luxembourgish as HL encounter more difficulties (Luxembourg Centre for Educational Testing (LUCET) & Service de Coordination de la Recherche et de l'Innovation pédagogiques et technologiques (SCRIPT), 2021). It is important to note for future research on this particular population, that the school curriculum underwent recent changes where French is now introduced in preschools (*The Luxembourgish Education System – An Overview*, n.d.).

12 Implications and future research

In the studies presented in this thesis, we focused on the verbal representation of numbers rather than on arithmetic and mathematical problem-solving. It is however assumed that the cognitive mechanisms discussed above are commonly used for solving arithmetic. This assumption is sustained by Study 2 that shows that the reaction times in the auditory-visual number-matching task correlate with the number of resolved arithmetic problems (see also Steiner, Banfi, et al., 2021 for similar results). Indeed, worse performances for solving arithmetic in French than in German have been found in previous research on the same population (Van Rinsveld, Schiltz, Brunner, et al., 2016). Hence LM2 cost in arithmetic could be the result of the multiple factors affecting bilingual numerical representation described above. Arithmetic problems can be resolved by direct fact retrieval in long-term memory or by algorithmic resolution. In the case of direct arithmetic fact retrieval, we can suppose similar mechanisms than for the LM2 lexical cost described above (i.e. 6 x 7 → 42). In the case of algorithmic resolution (i.e. 6 x 7 → 7 x 3 x 2 → 21 x 2 → 42) the cognitive mechanisms underpinning LM2 cost might add up from the multiple steps and could therefore be related to verbal working memory (Baddeley et al., 1975). Hence future research should focus on the effect of bilingualism on working memory.

To overcome the LM2 cost it is possible that bilinguals develop specific strategies to solve arithmetic such as translations or relying more heavily on visuo-spatial number representations. For translation, it could be that the arithmetic problem in the LM2 is mentally resolved in the most dominant language (i.e. the LM1) and then translating the solution in the LM2, implying a LSC. Another possibility is that bilinguals would rely more heavily on language-independent visuospatial numerical representations (see Van Rinsveld et al., 2017). Different strategies would also be interesting to compare on a working memory perspective in future studies.

Future investigations on working memory would be extremely important from a bilingual curriculum student’s perspective. This since the cognitive cost of processing numbers in a second language might be easier to overcome for high-achieving than low-achieving students. Low-achieving students might have fewer cognitive resources available to overcome the costs elicited by bilingual education (see for example McClung and Arya, (2018)). Hence a bilingual education system might be particularly detrimental to low-achieving students who would already struggle with a monolingual school curriculum. Ideally, future studies would need to be longitudinal with the challenge of tracking low-achieving students who are at risk of class repetition. Verbal working memory and long-term memory might be particularly impacted by the LM2 cost and would therefore be an interesting candidate to investigate the impact of bilingual education on high and low-achieving students.

An important point that would require more research regards the effect of resources and individual differences with regards to bilingual education. Luxembourg provides an ideal context to investigate the question of multilingualism. It is as having a “multilingualism ecological laboratory” to answer the challenges that come with it. For example one of the teacher resources that has been identified as a particular strength is teacher students relationships (see Emslander, 2024). We have seen that languages play a key role in mathematics and hence likely education. For this education context it would therefore be important for a bilingual curriculum to first establish second language proficiency and then teaching contents can be vehiculated through this language.

Part of the problem of assessing the role of AoA in L2 proficiency is the measure of language proficiency, which is often measured using self-rating scales (see Tomoschuk et al., 2019). Retrospective questionnaires are well known for not being very reliable. Many bilingual studies are based on those questionnaires or on assumptions based on characteristics of the sampled population (i.e. the education system). Hence, on a methodological level, future research should focus on finding an objective way to measure and compare language proficiencies in bilinguals. An objective measure of proficiency would therefore be important for bilingual research. Finally regarding lexico-semantic differences in bilingual number representations, it would be interesting to compare them for other semantic aspects of numbers such as for example ordinality, magnitude and parity.

Despite the cognitive costs reviewed before, there are undeniable major benefits of bilingualism that span beyond cognitive aspects. One of these benefits is the ability to communicate with people from different cultures and countries. Also knowing multiple languages allows one to immerse into a different culture, with direct access through literature, music, theatre, or popular expressions. Other benefits are enhanced employability and mobility. Hence the global cost-benefit evaluation of bilingualism needs to be done carefully. In part because the benefits listed above are harder to quantify than the cognitive costs. Ultimately, the cost-benefit perception of bilingualism might depend on individual, societal, and governmental priorities, and expectations. These perceptions are however important to enlighten with rigorous research on bilingualism since particularly bilingual school curriculums affect students' outcomes, in particular the ones with socio-economic and cognitive vulnerabilities.

13 General summary

In this thesis, I have reviewed and investigated bilingual lexical and semantic number representations. In three different studies, we compared German and French lexical representations, the influence of language-dependent morpho-syntax on accessing lexical representation and the associations with lexico-semantic representations. The bilingual samples in all three studies are Luxembourgish German-French bilinguals. In the Luxembourgish education system, German is the general instruction language for the first 6 years (i.e. LM1). Then the language for math instruction switches to French for the next 4 years (i.e. LM2) and gradually becomes the general instruction. In study 1 we found an LM2 cost in lexical retrieval for the LM2 compared to the LM1, (i.e. slower number naming). In addition to the LM2 cost, French '70s to '90s base-20 numerals were slower to process, indicating an effect of morpho-syntax on accessing lexical representations of numbers. This independently from increasing LM2 proficiency, the results replicate on four age groups with increasing proficiency. In study 2 we investigated the morpho-syntactic effect on German-French bilinguals and language-matched monolinguals. In an auditory-visual number matching task, we manipulated the visual presentation of two-digit numbers, either mimicking LM1's inverted unit-ten morpho-syntax or LM2's more transparent ten-unit morpho-syntax. We found that only the LM2 was affected by the morpho-syntactic experimental manipulation Moreover, we found a bilingual lexical cost such that bilingual in the LM1 German were slower than German monolinguals. In study 3 we compared lexico-semantic associations in bilinguals with a priming distance effect paradigm. The results indicate that while both language's equivalent primes facilitate number naming (i.e. "cinq" and "fünf" facilitate the naming of 5) the priming distance effect was only observed with LM1 primes. Since the priming distance effect arises from the association between different numerals, it is interpreted as weaker lexico-semantic associations of the LM2 compared to LM1. In a nutshell, this thesis presents empirical evidence for LM2 weaker lexical associations which are more impacted by linguistic morphos-syntax as well as weaker lexico-semantic associations compared to the LM1.

14 REFERENCES

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*, *128*(3), 466–478. https://doi.org/10.1016/j.actpsy.2008.03.014

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., Cappa, S. F., & Perani, D. (2001). The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition*, *4*(2), 179–190. https://doi.org/10.1017/S136672890100027X

Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., Cappa, S. F., & Costa, A. (2012). Bilingualism Tunes the Anterior Cingulate Cortex for Conflict Monitoring. *Cerebral Cortex*, *22*(9), 2076–2086. https://doi.org/10.1093/cercor/bhr287

Adesope, O. 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*(2), 207–245. https://doi.org/10.3102/0034654310368803

Almoammer, A., Sullivan, J., Donlan, C., Marusic, F., Zaucer, R., O'Donnell, T., & Barner, D. (2013). Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences*, *110*(46), 18448–18453. https://doi.org/10.1073/pnas.1313652110

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review*, *2*(3), 213–236. https://doi.org/10.1016/0273-2297(82)90012-0

Atagi, N., & Sandhofer, C. M. (2023). Monolingual and bilingual children's performance on arithmetic fluency varies by language fluency. *Journal of Experimental Child Psychology*, *233*, 105695. https://doi.org/10.1016/j.jecp.2023.105695

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 575–589. https://doi.org/10.1016/S0022-5371(75)80045-4

Bahnmueller, J., Göbel, S. M., Pixner, S., Dresen, V., & Moeller, K. (2020). More than simple facts: Cross-linguistic differences in place-value processing in arithmetic fact retrieval. *Psychological Research*, *84*(3), 650–659. https://doi.org/10.1007/s00426-018-1083-7

Bahnmueller, J., Moeller, K., Mann, A., & Nuerk, H.-C. (2015). On the limits of language influences on numerical cognition – no inversion effects in three-digit number magnitude processing in adults. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01216

Bahnmueller, J., Nuerk, H.-C., & Moeller, K. (2018). A Taxonomy Proposal for Types of Interactions of Language and Place-Value Processing in Multi-Digit Numbers. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01024

Baker, C., & Wright, W. E. (2021). *Foundations of Bilingual Education and Bilingualism*. Multilingual Matters. http://ebookcentral.proquest.com/lib/unilu-ebooks/detail.action?docID=6940556

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00328

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barrouillet, P., Camos, V., Perruchet, P., & Seron, X. (2004). ADAPT: A Developmental, Asemantic, and Procedural Model for Transcoding From Verbal to Arabic Numerals. *Psychological Review*, *111*(2), 368–394. https://doi.org/10.1037/0033-295X.111.2.368

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*(3), 201–221. https://doi.org/10.1016/S0010-0277(02)00178-6

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baumert, J., & Schümer, G. (2002). Family Background, Selection and Achievement: The German Experience. *Improving Schools*, *5*(3), 13–20. https://doi.org/10.1177/136548020200500304

Beal, C. R., Adams, N. M., & Cohen, P. R. (2010). Reading Proficiency and Mathematics Problem Solving by High School English Language Learners. *Urban Education*, *45*(1), 58–74. https://doi.org/10.1177/0042085909352143

Benni, S. (1996). *Terra!* Feltrinelli.

Benoit, L., Lehalle, H., Molina, M., Tijus, C., & Jouen, F. (2013). Young children's mapping between arrays, number words, and digits. *Cognition*, *129*(1), 95–101. https://doi.org/10.1016/j.cognition.2013.06.005

Bernardo, A. B. I. (2001). Asymmetric activation of number codes in bilinguals: Further evidence for the encoding complex model of number processing. *Memory & Cognition*, *29*(7), 968–976. https://doi.org/10.3758/BF03195759

Bernardo, A. B. I. (2002). Language and mathematical problem solving among bilinguals. *The Journal of Psychology*, *136*(3), 283–297. https://doi.org/10.1080/00223980209604156

Bernardo, A. B. I. (2005). Language and Modeling Word Problems in Mathematics Among Bilinguals. *The Journal of Psychology*, *139*(5), 413–425. https://doi.org/10.3200/JRLP.139.5.413-425

Bernardo, A. B. I., & Calleja, M. O. (2005). The Effects of Stating Problems in Bilingual Students' First and Second Languages on Solving Mathematical Word Problems. *The Journal of Genetic Psychology*, *166*(1), 117–129. https://doi.org/10.3200/GNTP.166.1.117-129

Bhatia, T. K., & Ritchie, W. C. (2013). *The handbook of bilingualism and multilingualism* (2nd ed.). Wiley-Blackwell.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*(1), 3–11. https://doi.org/10.1017/S1366728908003477

Bialystok, E., Craik, F. I. M., 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., & Feng, X. (2009). Language proficiency and executive control in proactive interference: Evidence from monolingual and bilingual children and adults. *Brain and Language*, *109*(2–3), 93–100. https://doi.org/10.1016/j.bandl.2008.09.001

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism (Cambridge, England)*, *13*(4), 525–531. https://doi.org/10.1017/S1366728909990423

Blanken, G., Dorn, M., & Sinn, H. (1997). Inversion Errors in Arabic Number Reading: Is There a Nonsemantic Route? *Brain and Cognition*, *34*(3), 404–423. https://doi.org/10.1006/brcg.1997.0917

Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (Pan troglodytes). *Journal of Comparative Psychology*, *103*(1), 23–31. https://doi.org/10.1037/0735-7036.103.1.23

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the Numerosities 1 to 9 by Monkeys. *Science*, *282*(5389), 746–749. https://doi.org/10.1126/science.282.5389.746

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General*, *124*(4), 434–452. https://doi.org/10.1037/0096-3445.124.4.434

Brysbaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism: Language and Cognition*, *13*(3), 359–371. https://doi.org/10.1017/S1366728909990344

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives*, *8*(1), 36–46. https://doi.org/10.1111/cdep.12059

Bürkner, P.-C. (2017). **brms**: An *R* Package for Bayesian Multilevel Models Using *Stan*. *Journal of Statistical Software*, *80*(1). https://doi.org/10.18637/jss.v080.i01

Bylund, E., Antfolk, J., Abrahamsson, N., Olstad, A. M. H., Norrman, G., & Lehtonen, M. (2022). Does bilingualism come with linguistic costs? A meta-analytic review of the bilingual lexical deficit. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-022-02136-7

Bylund, E., Antfolk, J., Abrahamsson, N., Olstad, A. M. H., Norrman, G., & Lehtonen, M. (2023). Does bilingualism come with linguistic costs? A meta-analytic review of the bilingual lexical deficit. *Psychonomic Bulletin & Review*, *30*(3), 897–913. https://doi.org/10.3758/s13423-022-02136-7

Camos, V. (2008). Low working memory capacity impedes both efficiency and learning of number transcoding in children. *Journal of Experimental Child Psychology*, *99*(1), 37–57. https://doi.org/10.1016/j.jecp.2007.06.006

Campbell, J. I. D. (1995). Mechanisms of Simple Addition and Multiplication: A Modified Network-interference Theory and Simulation. *Mathematical Cognition*, *1*, 121–164.

Campbell, J. I. D. (2005). Asymmetrical language switching costs in Chinese–English bilinguals' number naming and simple arithmetic. *Bilingualism: Language and Cognition*, *8*(1), 85–91. https://doi.org/10.1017/S136672890400207X

Campbell, J. I. D., & Clark, J. M. (1988). An encoding-complex view of cognitive number processing: Comment on McCloskey, Sokol, and Goodman (1986). *Journal of Experimental Psychology: General*, *117*(2), 204–214. https://doi.org/10.1037/0096-3445.117.2.204

Campbell, J. I. D., & Clark, J. M. (1992). Chapter 12 Cognitive Number Processing: An Encoding-Complex Perspective. In J. I. D. Campbell (Ed.), *Advances in Psychology* (Vol. 91, pp. 457–491). North-Holland. https://doi.org/10.1016/S0166-4115(08)60894-8

Campbell, J. I. D., & Epp, L. J. (2004). An Encoding-Complex Approach to Numerical Cognition in Chinese-English Bilinguals. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *58*(4), 229–244. https://doi.org/10.1037/h0087447

Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General*, *130*(2), 299–315. https://doi.org/10.1037/0096-3445.130.2.299

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Carey, S., & Barner, D. (2019). Ontogenetic Origins of Human Integer Representations. *Trends in Cognitive Sciences*, *23*(10), 823–835. https://doi.org/10.1016/j.tics.2019.07.004

Carey, S., Shusterman, A., Haward, P., & Distefano, R. (2017). Do analog number representations underlie the meanings of young children's verbal numerals? *Cognition*, *168*, 243–255. https://doi.org/10.1016/j.cognition.2017.06.022

Cerda, V. R., Grenier, A. E., & Wicha, N. Y. Y. (2019). Bilingual children access multiplication facts from semantic memory equivalently across languages: Evidence from the N400. *Brain and Language*, *198*, 104679. https://doi.org/10.1016/j.bandl.2019.104679

Chincotta, D., & Underwood, G. (1996). Mother tongue, language of schooling and bilingual digit span. *British Journal of Psychology*, *87*(2), 193–208. https://doi.org/10.1111/j.2044-8295.1996.tb02585.x

Chincotta, D., & Underwood, G. (1997). Speech Rate Estimates, Language of Schooling and Bilingual Digit Span. *European Journal of Cognitive Psychology*, *9*(3), 325–348. https://doi.org/10.1080/713752562

Chow, J. C., Majeika, C. E., & Sheaffer, A. W. (2021). Language skills of children with and without mathematics difficulty. *Journal of Speech, Language & Hearing Research*, *64*(9), 3571–3577. https://doi.org/10.1044/2021_JSLHR-20-00378

Chrisomalis, S. (2010). *Numerical Notation: A Comparative History*. Cambridge University Press.

Clayton, F. J., Copper, C., Steiner, A. F., Banfi, C., Finke, S., Landerl, K., & Göbel, S. M. (2020). Two-digit number writing and arithmetic in Year 1 children: Does number word inversion matter? *Cognitive Development*, *56*, 100967. https://doi.org/10.1016/j.cogdev.2020.100967

Clement, D. H., Sarama, J., & Germeroth, C. (2016). Learning executive function and early mathematics: Directions of causal relations. *Early Childhood Research Quarterly*, *36*, 79–90.

Cohen, J., Mac Whinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers*, *25*(2), 257–271.

Cohen, L., & Dehaene, S. (1991). Neglect Dyslexia for Numbers? A Case Report. *Cognitive Neuropsychology*, *8*(1), 39–58. https://doi.org/10.1080/02643299108253366

Colomé, àngels, Laka, I., & Sebastián-Gallés, N. (2010). Language effects in addition: How you say it counts. *Quarterly Journal of Experimental Psychology*, *63*(5), 965–983. https://doi.org/10.1080/17470210903134377

Colomé, À. (2001). Lexical Activation in Bilinguals' Speech Production: Language-Specific or Language-Independent? *Journal of Memory and Language*, *45*(4), 721–736. https://doi.org/10.1006/jmla.2001.2793

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*(4), 589–608. https://doi.org/10.1037/0033-295X.100.4.589

Comrie, B. (2013). Numeral bases (v2020.3). In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*. Zenodo. https://doi.org/10.5281/zenodo.7385533

Contreras-Saavedra, C. E., Koch, I., Schuch, S., & Philipp, A. M. (2021). The reliability of language-switch costs in bilingual one- and two-digit number naming. *International Journal of Bilingualism*, *25*(1), 272–285. https://doi.org/10.1177/1367006920951873

Contreras-Saavedra, C. E., Willmes, K., Koch, I., Schuch, S., Benini, E., & Philipp, A. M. (2020). Multilingual two-digit number naming: The influence of composition rules on language switching. *Quarterly Journal of Experimental Psychology*, 1747021820916108. https://doi.org/10.1177/1747021820916108

Coolen, I., Merkley, R., Ansari, D., Dove, E., Dowker, A., Mills, A., Murphy, V., von Spreckelsen, M., & Scerif, G. (2021). Domain-general and domain-specific influences on emerging numerical cognition: Contrasting uni-and bidirectional prediction models. *Cognition*, *215*, 104816. https://doi.org/10.1016/j.cognition.2021.104816

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002

Cummins, J. (2016). *L'éducation bilingue: Perspectives internationales sur la recherche et les politiques linguistiques éducatives. L'éducation bilingue en France. Politiques linguistiques, modèles et pratiques* (pp. 529–244). Lambert-Lucas.

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

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science*, *26*(1), 99–107. https://doi.org/10.1177/0956797614557866

de Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. Psychology Press. https://doi.org/10.4324/9780203841228

De Visscher, A., & Noël, M.-P. (2014). The detrimental effect of interference in multiplication facts storing: Typical development and individual differences. *Journal of Experimental Psychology: General*, *143*(6), 2380–2400. https://doi.org/10.1037/xge0000029

De Vos, T. (1992). Tempo test rekenen (TTR)[Arithmetic number fact test]. *Nijmegen: Berkhout*.

Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review*, *22*(6), 1630–1645. https://doi.org/10.3758/s13423-015-0836-1

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Rev. and updated ed). Oxford University Press.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The Mental Representation of Parity and Number Magnitude. *Journal of Experimental Psychology*, *122*(3), 371–396.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, *1*(1), 83–120.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*(1), 1–29. https://doi.org/10.1016/0010-0277(92)90030-l

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*(3–6), 487–506. https://doi.org/10.1080/02643290244000239

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence. *Science*, *284*(5416), 970–974. https://doi.org/10.1126/science.284.5416.970

Del Maschio, N., & Abutalebi, J. (2019). Language Organization in the Bilingual and Multilingual Brain. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the Neuroscience of Multilingualism* (pp. 197–213). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119387725.ch9

Delazer, M., & Benke, T. (1997). Arithmetic Facts without Meaning. *Cortex*. https://doi.org/10.1016/S0010-9452(08)70727-5

Deloche, G., & Seron, X. (1982). From one to 1: An analysis of a transcoding process by means of neuropsychological data. *Cognition*, *12*(2), 119–149. https://doi.org/10.1016/0010-0277(82)90009-9

den Heyer, K., & Briand, K. (1986). Priming Single Digit Numbers: Automatic Spreading Activation Dissipates as a Function of Semantic Distance. *The American Journal of Psychology*, *99*(3), 315. https://doi.org/10.2307/1422488

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, *82*(1), 64–81. https://doi.org/10.1348/2044-8279.002002

Dewaele, J.-M. (2007). Multilinguals' language choice for mental calculation. *Intercultural Pragmatics*, *4*(3). https://doi.org/10.1515/IP.2007.017

Dewaele, J.-M., Botes, E., & Meftah, R. (2023). A Three-Body Problem: The effects of foreign language anxiety, enjoyment, and boredom on academic achievement. *Annual Review of Applied Linguistics*, *43*, 7–22. https://doi.org/10.1017/S0267190523000016

Dick, A. S., Garcia, N. L., Pruden, S. M., Thompson, W. K., Hawes, S. W., Sutherland, M. T., Riedel, M. C., Laird, A. R., & Gonzalez, R. (2019). No evidence for a bilingual executive function advantage in the ABCD study. *Nature Human Behaviour*, *3*(7), 692–701. https://doi.org/10.1038/s41562-019-0609-3

Dijkstra, T. (2005). Bilingual Visual Word Recognition and Lexical Access. In *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). Oxford University Press. https://doi.org/10.1017/S0272263107210071

Dijkstra, T., & Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In *Localist connectionist approaches to human cognition* (pp. 189–225). Lawrence Erlbaum Associates Publishers.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Wahl, A., Buytenhuijs, F., Halem, N. V., Al-Jibouri, Z., Korte, M. D., & Rekké, S. (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Dotan, D. (2023). Top-Down Number Reading: Language Affects the Visual Identification of Digit Strings. *Cognitive Science*, *47*(10), e13368. https://doi.org/10.1111/cogs.13368

Dotan, D., & Friedmann, N. (2018). A cognitive model for multidigit number reading: Inferences from individuals with selective impairments. *Cortex*, *101*, 249–281. https://doi.org/10.1016/j.cortex.2017.10.025

Dowker, A., Bala, S., & Lloyd, D. (2008). Linguistic Influences on Mathematical Development: How Important Is the Transparency of the Counting System? *Philosophical Psychology*, *21*(4), 523–538. https://doi.org/10.1080/09515080802285511

Dowker, A., & Nuerk, H.-C. (2016). Editorial: Linguistic Influences on Mathematics. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01035

Dowker, A., & Roberts, M. (2015). Does the transparency of the counting system affect children's numerical abilities? *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00945

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

Duyck, W., & Brysbaert, M. (2002). What number translation studies can teach us about the lexico-semantic organisation in bilinguals. *Psychologica Belgica*, *42*(3), 151–175.

Duyck, W., & Brysbaert, M. (2004). Forward and Backward Number Translation Requires Conceptual Mediation in Both Balanced and Unbalanced Bilinguals. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(5), 889–906. https://doi.org/10.1037/0096-1523.30.5.889

Duyck, W., & Brysbaert, M. (2008). Semantic Access in Number Word Translation: The Role of Crosslingual Lexical Similarity. *Experimental Psychology*, *55*(2), 102–112. https://doi.org/10.1027/1618-3169.55.2.102

Duyck, W., Depestel, I., Fias, W., & Reynvoet, B. (2008). Cross-lingual numerical distance priming with second-language number words in native- to third-language number word translation. *Quarterly Journal of Experimental Psychology*, *61*(9), 1281–1290. https://doi.org/10.1080/17470210802000679

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of Acquisition Effects in Adult Lexical Processing Reflect Loss of Plasticity in Maturing Systems: Insights From Connectionist Networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1103–1123. https://doi.org/10.1037/0278-7393.26.5.1103

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology*, *71*(1), 43–51. https://doi.org/10.1111/j.2044-8295.1980.tb02728.x

Emslander, V. (2024). *An Exploration of Factors Driving School Success in Diverse Students Through Meta-Analytic and Value-Added Modeling*. Luxembourg.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Ester, P., Morales, I., Moraleda, Á., & Bermejo, V. (2021). The verbal component of mathematical problem solving in bilingual contexts by early elementary schoolers. *Mathematics*, *9*(5), 5. https://doi.org/10.3390/math9050564

Fabbro, F., Skrap, M., & Aglioti, S. (2000). Pathological switching between languages after frontal lesions in a bilingual patient. *Journal of Neurology, Neurosurgery & Psychiatry*, *68*(5), 650–652. https://doi.org/10.1136/jnnp.68.5.650

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Fernandes, A. (2023). Understanding mathematics preservice teachers' beliefs about English learners using the Mathematics Education for English Learners Scale (MEELS). *Bilingual Research Journal*, *45*(3–4), 295–314. https://doi.org/10.1080/15235882.2023.2169406

Fernández-Coello, A., Havas, V., Juncadella, M., Sierpowska, J., Rodríguez-Fornells, A., & Gabarrós, A. (2016). Age of language acquisition and cortical language organization in multilingual patients undergoing awake brain mapping. *Journal of Neurosurgery*, *126*(6), 1912–1923. https://doi.org/10.3171/2016.5.JNS152791

Fiebach, C. J., Friederici, A. D., Müller, K., von Cramon, D. Y., & Hernandez, A. E. (2003). Distinct brain representations for early and late learned words. *NeuroImage*, *19*(4), 1627–1637. https://doi.org/10.1016/S1053-8119(03)00227-1

Finger, H., Goeke, C., Diekamp, D., Standvoß, K., & König, P. (2017). LabVanced: A unified JavaScript framework for online studies. *International Conference on Computational Social Science (Cologne)*. https://www.labvanced.com/

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age Constraints on Second-Language Acquisition. *Journal of Memory and Language*, *41*(1), 78–104. https://doi.org/10.1006/jmla.1999.2638

Fodor, J. A. (1985). Précis of The Modularity of Mind. *THE BEHAVIORAL AND BRAIN SCIENCES*.

Forsyth, S. R., & Powell, S. R. (2017). Differences in the mathematics-vocabulary knowledge of fifth-grade students with and without learning difficulties. *Learning Disabilities Research & Practice*, *32*(4), 231–245. https://doi.org/10.1111/ldrp.12144

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*(3), 819–824. https://doi.org/10.1016/j.cognition.2008.04.007

French, R. M., & Jacquet, M. (2004). Understanding bilingual memory: Models and data. *Trends in Cognitive Sciences*, *8*(2), 87–93. https://doi.org/10.1016/j.tics.2003.12.011

Frenck-Mestre, C., & Vaid, J. (1993). Activation of number facts in bilinguals. *Memory & Cognition*, *21*(6), 809–818. https://doi.org/10.3758/BF03202748

Friederici, A. D. (2017). *Language in our brain: The origins of a uniquely human capacity*. The MIT Press.

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. https://doi.org/10.1016/j.edurev.2013.05.003

Fuchs, L. S., Fuchs, D., Compton, D. L., Hamlett, C. L., & Wang, A. Y. (2015). Is word-problem solving a form of text comprehension? Scientific studies of reading. *The Official Journal of the Society for the Scientific Study of Reading*, *19*(3), 204–223. https://doi.org/10.1080/10888438.2015.1005745

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, *98*(1), 29–43. https://doi.org/10.1037/0022-0663.98.1.29

Garcia, O., Faghihi, N., Raola, A. R., & Vaid, J. (2021). Factors influencing bilinguals' speed and accuracy of number judgments across languages: A meta-analytic review. *Journal of Memory and Language*, *118*, 104211. https://doi.org/10.1016/j.jml.2020.104211

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, *8*(3), 259. https://doi.org/10.1007/BF03174081

Geary, D. C., Berch, D. B., & Koepke, K. M. (Eds.). (2015). *Evolutionary origins and early development of number processing* (First edition). Academic Press.

Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of Arithmetical Competencies in Chinese and American Children: Influence of Age, Language, and Schooling. *Child Development*, *67*(5), 2022–2044. https://doi.org/10.1111/j.1467-8624.1996.tb01841.x

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' Functional Numeracy Is Predicted by Their School Entry Number System Knowledge. *PLOS ONE*, *8*(1), e54651. https://doi.org/10.1371/journal.pone.0054651

Gelman, R., & Butterworth, B. (2005). Number and language: How are they related? *Trends in Cognitive Sciences*, *9*(1), 6–10. https://doi.org/10.1016/j.tics.2004.11.004

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Harvard University Press.

Gillet, S., Barbu, C., & Poncelet, M. (2021). Early bilingual immersion school program and cognitive development in French-speaking children: Effect of the second language learned (English vs. Dutch) and exposition duration (2 vs. 5 years). *PLOS ONE*, *16*(10), e0258458. https://doi.org/10.1371/journal.pone.0258458

Göbel, S. M., Moeller, K., Pixner, S., Kaufmann, L., & Nuerk, H.-C. (2014). Language affects symbolic arithmetic in children: The case of number word inversion. *Journal of Experimental Child Psychology*, *119*, 17–25. https://doi.org/10.1016/j.jecp.2013.10.001

Göbel, S. M., Watson, S. E., Lervåg, A., & Hulme, C. (2014). Children's Arithmetic Development: It Is Number Knowledge, Not the Approximate Number Sense, That Counts. *Psychological Science*, *25*(3), 789–798. https://doi.org/10.1177/0956797613516471

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., Herman, D. H., Clasen, L. S., Toga, A. W., & others. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(21), 8174–8179.

Gold, B. T., Johnson, N. F., & Powell, D. K. (2013). Lifelong bilingualism contributes to cognitive reserve against white matter integrity declines in aging. *Neuropsychologia*, *51*(13), 2841–2846. https://doi.org/10.1016/j.neuropsychologia.2013.09.037

Goldsmith, S. F., El-Baba, M., He, X., Lewis, D. J., Akoury Dirani, L., Liu, J., & Morton, J. B. (2023). No bilingual advantage in children's attentional disengagement: Congruency and sequential congruency effects in a large sample of monolingual and bilingual children. *Journal of Experimental Child Psychology*, *233*, 105692. https://doi.org/10.1016/j.jecp.2023.105692

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis☆. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Gollan, T. H., Sandoval, T., & Salmon, D. P. (2011). Cross-Language Intrusion Errors in Aging Bilinguals Reveal the Link Between Executive Control and Language Selection. *Psychological Science*, *22*(9), 1155–1164. https://doi.org/10.1177/0956797611417002

Gordon, P. (2004). Numerical Cognition Without Words: Evidence from Amazonia. *Science*, *306*(5695), 496–499. https://doi.org/10.1126/science.1094492

Grabner, R. H., Saalbach, H., & Eckstein, D. (2012a). Language-Switching Costs in Bilingual Mathematics Learning. *Mind, Brain, and Education*, *6*(3), 147–155. https://doi.org/10.1111/j.1751-228X.2012.01150.x

Grabner, R. H., Saalbach, H., & Eckstein, D. (2012b). Language-Switching Costs in Bilingual Mathematics Learning. *Mind, Brain, and Education*, *6*(3), 147–155. https://doi.org/10.1111/j.1751-228X.2012.01150.x

Grainger, J., Midgley, K., & Holcomb, P. J. (2010). Chapter 14. Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In M. Kail & M. Hickmann (Eds.), *Language Acquisition and Language Disorders* (Vol. 52, pp. 267–283). John Benjamins Publishing Company. https://doi.org/10.1075/lald.52.18gra

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

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Greer, B., Verschaffel, L., & Corte, E. (2002). The answer is really 4.5": Beliefs about word problems. In G. C. Leder, E. Pehkonen, & G. Törner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 271–292). Springer Netherlands. https://doi.org/10.1007/0-306-47958-3_16

Greisen, M., Georges, C., Hornung, C., Sonnleitner, P., & Schiltz, C. (2021). Learning mathematics with shackles: How lower reading comprehension in the language of mathematics instruction accounts for lower mathematics achievement in speakers of different home languages. *Acta Psychologica*, *221*, 103456. https://doi.org/10.1016/j.actpsy.2021.103456

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3–15. https://doi.org/10.1016/0093-934X(89)90048-5

Grosjean, F. (2001). *The Bilingual's Language Modes*. Blackwell Publishing.

Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press.

Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*(3), 325–340. https://doi.org/10.1177/0267658316678286

Gunderson, E. A., & Levine, S. C. (2011). Some types of parent number talk count more than others: Relations between parents' input and children's cardinal-number knowledge. *Developmental Science*, *14*(5), 1021–1032. https://doi.org/10.1111/j.1467-7687.2011.01050.x

Gunnerud, H. L., Ten Braak, D., Reikerås, E. K. L., Donolato, E., & Melby-Lervåg, M. (2020). Is bilingualism related to a cognitive advantage in children? A systematic review and meta-analysis. *Psychological Bulletin*, *146*(12), 1059–1083. https://doi.org/10.1037/bul0000301

Hahn, C. G. K., Saalbach, H., & Grabner, R. H. (2017). Language-dependent knowledge acquisition: Investigating bilingual arithmetic learning. *Bilingualism: Language and Cognition*, *22*(1), 47–57. https://doi.org/10.1017/S1366728917000530

Hahn, C. G. K., Saalbach, H., & Grabner, R. H. (2019). Language-dependent knowledge acquisition: Investigating bilingual arithmetic learning. *Bilingualism: Language and Cognition*, *22*(1), 47–57. https://doi.org/10.1017/S1366728917000530

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, *109*(28), 11116–11120. https://doi.org/10.1073/pnas.1200196109

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). *Individual differences in non-verbal number acuity correlate with maths achievement*. *455*, 655–668. https://doi.org/10.1038/nature07246

Hartanto, A., Yang, H., & Yang, S. (2018). Bilingualism positively predicts mathematical competence: Evidence from two large-scale studies. *Learning and Individual Differences*, *61*, 216–227. https://doi.org/10.1016/j.lindif.2017.12.007

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, *177*, 263–277. https://doi.org/10.1016/j.cognition.2018.04.007

Haspelmath, M., Dryer, M. S., Gil, D., & Comrie, B. (2005). *The World Atlas of Language Structures*. Oxford Univ. Press.

Heppt, B., Haag, N., Böhme, K., & Stanat, P. (2015). The Role of Academic-Language Features for Reading Comprehension of Language-Minority Students and Students From Low-SES Families. *Reading Research Quarterly*, *50*(1), 61–82. https://doi.org/10.1002/rrq.83

Hernandez, A. E. (2013). *The bilingual brain*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199828111.001.0001

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, *18*(4), 625–658. https://doi.org/10.3758/s13423-011-0116-7

Hirsh, K. W., Morrison, C. M., Gaset, S., & Carnicer, E. (2003). Age of acquisition and speech production in L2. *Bilingualism: Language and Cognition*, *6*(2), 117–128. https://doi.org/10.1017/S136672890300107X

Ho, C. S.-H., & Fuson, K. C. (1998). Children's Knowledge of Teen Quantities as Tens and Ones: Comparisons of Chinese, British, and American Kindergartners. *Journal of Educational Psychology*.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(1), 1–27. https://doi.org/10.1017/S0305000910000759

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, *103*(1), 17–29. https://doi.org/10.1016/j.jecp.2008.04.001

Hosoda, C., Tanaka, K., Nariai, T., Honda, M., & Hanakawa, T. (2013). Dynamic Neural Network Reorganization Associated with Second Language Vocabulary Acquisition: A Multimodal Imaging Study. *Journal of Neuroscience*, *33*(34), 13663–13672. https://doi.org/10.1523/JNEUROSCI.0410-13.2013

Howard, S. R., Avarguès-Weber, A., Garcia, J. E., Greentree, A. D., & Dyer, A. G. (2018). Numerical ordering of zero in honey bees. *Science*, *360*(6393), 1124–1126. https://doi.org/10.1126/science.aar4975

Howard, S. R., Avarguès-Weber, A., Garcia, J. E., Greentree, A. D., & Dyer, A. G. (2019). Symbolic representation of numerosity by honeybees (Apis mellifera): Matching characters to small quantities. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1904), 20190238. https://doi.org/10.1098/rspb.2019.0238

Hyde, D. C. (2011). Two systems of non-symbolic numerical cognition. *Frontiers in Human Neuroscience*, *5*(NOVEMBER), 1–8. https://doi.org/10.3389/fnhum.2011.00150

Ifrah, G., & Bellos, D. (2000). *The universal history of numbers: From prehistory to the invention of the computer*. Wiley.

Imbo, I., Vanden Bulcke, C., De Brauwer, J., & Fias, W. (2014). Sixty-four or four-and-sixty? The influence of language and working memory on children's number transcoding. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00313

Indefrey, P. (2006). A Meta-analysis of Hemodynamic Studies on First and Second Language Processing: Which Suggested Differences Can We Trust and What Do They Mean? *Language Learning*, *56*(s1), 279–304. https://doi.org/10.1111/j.1467-9922.2006.00365.x

Itard, J.-M.-G. (1774-1838) A. du texte, Delasiauve, L. (1804-1893) A. du texte, Bousquet, J. B. É. (1794-1872) A. du texte, & Bousquet, J. B. É. (1794-1872) A. du texte. (1891). *Rapports et mémoires sur le sauvage de l'Aveyron, l'idiotie et la surdi-mutité / par Itard ; préface par Bournevillepar A. Bousquet*. https://gallica.bnf.fr/ark:/12148/bpt6k5752734w

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, *127*(2), 277–288. https://doi.org/10.1016/j.actpsy.2007.06.003

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition*, *106*(3), 1221–1247. https://doi.org/10.1016/j.cognition.2007.06.004

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106

Jackson, G. M., Swainson, R., Cunnington, R., & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, *4*(2), 169–178. https://doi.org/10.1017/S1366728901000268

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0

Johnson, M., & Munakata, Y. (2005). Processes of change in brain and cognitive development. *Trends in Cognitive Sciences*, *9*(3), 152–158. https://doi.org/10.1016/j.tics.2005.01.009

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The Discrimination of Visual Number. *The American Journal of Psychology*, *62*(4), 498. https://doi.org/10.2307/1418556

Kempert, S., Saalbach, H., & Hardy, I. (2011). Cognitive benefits and costs of bilingualism in elementary school students: The case of mathematical word problems. *Journal of Educational Psychology*, *103*(3), 547–561. https://doi.org/10.1037/a0023619

Klaus, J., & Schriefers, H. (2019). Bilingual Word Production. In J. W. Schwieter & M. Paradis (Eds.), *The Handbook of the Neuroscience of Multilingualism* (pp. 214–229). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119387725.ch10

Kleemans, T., & Segers, E. (2020). Linguistic precursors of advanced math growth in first-language and second-language learners. *Research in Developmental Disabilities*, *103*, 103661. https://doi.org/10.1016/j.ridd.2020.103661

Kleemans, T., Segers, E., & Verhoeven, L. (2014). Cognitive and linguistic predictors of basic arithmetic skills: Evidence from first-language and second-language learners. *International Journal of Disability, Development and Education*, *61*(3), 306–316. https://doi.org/10.1080/1034912X.2014.934017

Kleemans, T., Segers, E., & Verhoeven, L. (2018). Role of linguistic skills in fifth-grade mathematics. *Journal of Experimental Child Psychology*, *167*, 404–413. https://doi.org/10.1016/j.jecp.2017.11.012

Klein, E., Bahnmueller, J., Mann, A., Pixner, S., Kaufmann, L., Nuerk, H.-C., & Moeller, K. (2013). Language influences on numerical development—Inversion effects on multi-digit number processing. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00480

Kochari, A. R. (2019). Conducting Web-Based Experiments for Numerical Cognition Research. *Journal of Cognition*, *2*(1), 39. https://doi.org/10.5334/joc.85

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed Numbers: Exploring the Modularity of Numerical Representations With Masked and Unmasked Semantic Priming. *Journal of Experimental Psychology: Human Perception and Performance*, 24. https://doi.org/10.1037/0096-1523.25.6.1882

Kolers, P. A. (1968). Bilingualism and Information Processing. *Scientific American*, *218*(3), 78–86. https://doi.org/10.1038/scientificamerican0368-78

Koponen, T., Salmi, P., Eklund, K., & Aro, T. (2013). Counting and RAN: Predictors of arithmetic calculation and reading fluency. *Journal of Educational Psychology*, *105*(1), 162–175. https://doi.org/10.1037/a0029285

Kovelman, I., Baker, S. A., & Petitto, L.-A. (2008). Age of first bilingual language exposure as a new window into bilingual reading development. *Bilingualism: Language and Cognition*, *11*(2), 203–223. https://doi.org/10.1017/S1366728908003386

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01795

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2018). Symbolic Number Comparison Is Not Processed by the Analog Number System: Different Symbolic and Non-symbolic Numerical Distance and Size Effects. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00124

Krajewski, K., & Schneider, W. (2009). Exploring the impact of phonological awareness, visual–spatial working memory, and preschool quantity–number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology*, *103*(4), 516–531. https://doi.org/10.1016/j.jecp.2009.03.009

Kraut, C., & Pixner, S. (2022). Bilingual adults practicing multiplication tables – looking into bilingual arithmetic learning. *International Journal of Bilingual Education and Bilingualism*, *25*(5), 1825–1837. https://doi.org/10.1080/13670050.2020.1810204

Krinzinger, H., Gregoire, J., Desoete, A., Kaufmann, L., Nuerk, H.-C., & Willmes, K. (2011). Differential Language Effects on Numerical Skills in Second Grade. *Journal of Cross-Cultural Psychology*, *42*(4), 614–629. https://doi.org/10.1177/0022022111406252

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

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language*, *33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, *13*(3), 373–381. https://doi.org/10.1017/S136672891000009X

Lachelin, R., Marinova, M., Reynvoet, B., & Schiltz, C. (2023). Weaker semantic priming effects with number words in the second language of math learning. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001526

Lachelin, R., Rinsveld, A. van, Poncin, A., & Schiltz, C. (2022). Number transcoding in bilinguals—A transversal developmental study. *PLOS ONE*, *17*(8), e0273391. https://doi.org/10.1371/journal.pone.0273391

Lafay, A., Adrien, E., Lonardo Burr, S. D., Douglas, H., Provost-Larocque, K., Xu, C., LeFevre, J.-A., Maloney, E. A., Osana, H. P., Skwarchuk, S.-L., & Wylie, J. (2023). Transcoding of French numbers for first- and second-language learners in third grade. *Quarterly Journal of Experimental Psychology*, *77*(2), 393–407. https://doi.org/10.1177/17470218231174339

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*(2), 395–438. https://doi.org/10.1016/j.cognition.2006.10.005

Le Corre, M., Li, P., Huang, B. H., Jia, G., & Carey, S. (2016). Numerical morphology supports early number word learning: Evidence from a comparison of young Mandarin and English learners. *Cognitive Psychology*, *88*, 162–186. https://doi.org/10.1016/j.cogpsych.2016.06.003

Lê, M.-L., & Noël, M.-P. (2021). Preschoolers' mastery of advanced counting: The best predictor of addition skills 2 years later. *Journal of Experimental Child Psychology*, *212*, 105252. https://doi.org/10.1016/j.jecp.2021.105252

Lê, M.-L. T., & Noël, M.-P. (2020). Transparent number-naming system gives only limited advantage for preschooler's numerical development: Comparisons of Vietnamese and French-speaking children. *PLOS ONE*, *15*(12), e0243472. https://doi.org/10.1371/journal.pone.0243472

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

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: Dissociable systems. *Neuropsychologia*, *41*(14), 1942–1958. https://doi.org/10.1016/S0028-3932(03)00123-4

Lenneberg, E. H. (1967). The Biological Foundations of Language. *Hospital Practice*, *2*(12), 59–67. https://doi.org/10.1080/21548331.1967.11707799

Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means* [Manual]. https://CRAN.R-project.org/package=emmeans

L'essentiel. (n.d.). Retrieved April 6, 2024, from https://www.lessentiel.lu/fr

Li, P., & Farkas, I. (2002). 3 A self-organizing connectionist model of bilingual processing. In R. R. Heredia & J. Altarriba (Eds.), *Advances in Psychology* (Vol. 134, pp. 59–85). North-Holland. https://doi.org/10.1016/S0166-4115(02)80006-1

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex*, *58*, 301–324. https://doi.org/10.1016/j.cortex.2014.05.001

Lin, J.-F. L., Imada, T., & Kuhl, P. K. (2012). Mental Addition in Bilinguals: An fMRI Study of Task-Related and Performance-Related Activation. *Cerebral Cortex*, *22*(8), 1851–1861. https://doi.org/10.1093/cercor/bhr263

Lin, J.-F. L., Imada, T., & Kuhl, P. K. (2019). Neuroplasticity, bilingualism, and mental mathematics: A behavior-MEG study. *Brain and Cognition*, *134*, 122–134. https://doi.org/10.1016/j.bandc.2019.03.006

Lipton, J. S., & Spelke, E. S. (2005). Preschool Children's Mapping of Number Words to Nonsymbolic Numerosities. *Child Development*, *76*(5), 978–988. https://doi.org/10.1111/j.1467-8624.2005.00891.x

Liu, H., & Cao, F. (2016). L1 and L2 processing in the bilingual brain: A meta-analysis of neuroimaging studies. *Brain and Language*, *159*, 60–73. https://doi.org/10.1016/j.bandl.2016.05.013

Lochy, A., Domahs, F., Bartha, L., & Delazer, M. (2004). Specific order impairment in arabic number writing: A case-study. *Cognitive Neuropsychology*, *21*(5), 555–575. https://doi.org/10.1080/02643290342000618

Lochy, A., & Schiltz, C. (2022). An input lexicon for familiar numbers. *Journal of Numerical Cognition*, *8*(2), 244–258. https://doi.org/10.5964/jnc.7385

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.

Loi du 24 février 1984 sur le régime des langues. - Legilux. (n.d.). Retrieved April 6, 2024, from https://legilux.public.lu/eli/etat/leg/loi/1984/02/24/n1/jo

Lonnemann, J., Li, S., Zhao, P., Linkersdörfer, J., Lindberg, S., Hasselhorn, M., & Yan, S. (2019). Differences in Counting Skills Between Chinese and German Children Are Accompanied by Differences in Processing of Approximate Numerical Magnitude Information. *Frontiers in Psychology*, *9*, 2656. https://doi.org/10.3389/fpsyg.2018.02656

Lonnemann, J., & Yan, S. (2015). Does number word inversion affect arithmetic processes in adults? *Trends in Neuroscience and Education*, *4*(1), 1–5. https://doi.org/10.1016/j.tine.2015.01.002

Lowe, C. J., Cho, I., Goldsmith, S. F., & Morton, J. B. (2021). The Bilingual Advantage in Children's Executive Functioning Is Not Related to Language Status: A Meta-Analytic Review. *Psychological Science*, *32*(7), 1115–1146. https://doi.org/10.1177/0956797621993108

Luk, G., Bialystok, E., Craik, F. I. M., & Grady, C. L. (2011). Lifelong Bilingualism Maintains White Matter Integrity in Older Adults. *The Journal of Neuroscience*, *31*(46), 16808–16813. https://doi.org/10.1523/JNEUROSCI.4563-11.2011

Luk, G., Green, D. W., Abutalebi, J., & Grady, C. (2012). Cognitive control for language switching in bilinguals: A quantitative meta-analysis of functional neuroimaging studies. *Language and Cognitive Processes*, *27*(10), 1479–1488. https://doi.org/10.1080/01690965.2011.613209

Luxembourg Centre for Educational Testing (LUCET) & Service de Coordination de la Recherche et de l'Innovation pédagogiques et technologiques (SCRIPT). (2021). *Rapport National sur l'Éducation au Luxembourg 2021*. Luxembourg Centre for Educational Testing (LUCET) & Service de Coordination de la Recherche et de l'Innovation pédagogiques et technologiques (SCRIPT). https://doi.org/10.48746/BB2021LU-FR-DIGIPUB

Luxemburger Wort. (n.d.). Retrieved April 6, 2024, from https://www.wort.lu/

Mägiste, E. (1979). The competing language systems of the multilingual: A developmental study of decoding and encoding processes. *Journal of Verbal Learning and Verbal Behavior*, *18*(1), 79–89. https://doi.org/10.1016/S0022-5371(79)90584-X

Major, C. S., Paul, J. M., & Reeve, R. A. (2017). TEMA and Dot Enumeration Profiles Predict Mental Addition Problem Solving Speed Longitudinally. *Frontiers in Psychology*, *8*. https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02263

Mancilla-Martinez, J., & Lesaux, N. K. (2010). Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology*, *102*(3), 701.

Marchand, E., Wade, S., Sullivan, J., & Barner, D. (2020). Language-specific numerical estimation in bilingual children. *Journal of Experimental Child Psychology*, *197*, 104860. https://doi.org/10.1016/j.jecp.2020.104860

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLOS ONE*, *7*(8). https://doi.org/10.1371/journal.pone.0043230

Marian, V., & Fausey, C. M. (2006). Language-dependent memory in bilingual learning. *Applied Cognitive Psychology*, *20*(8), 1025–1047. https://doi.org/10.1002/acp.1242

Marian, V., Shook, A., & Schroeder, S. R. (2013). Bilingual two-way immersion programs benefit academic achievement. *Bilingual Research Journal*, *36*(2), 167–186. https://doi.org/10.1080/15235882.2013.818075

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115. https://doi.org/10.1017/S1366728903001068

Marinova, M., Georges, C., Guillaume, M., Reynvoet, B., Schiltz, C., & Van Rinsveld, A. (2021). Automatic integration of numerical formats examined with frequency-tagged EEG. *Scientific Reports*, *11*(1), 21405. https://doi.org/10.1038/s41598-021-00738-0

Marsh, L. G., & Maki, R. H. (1976). Efficiency of arithmetic operations in bilinguals as a function of language. *Memory & Cognition*, *4*(4), 459–464. https://doi.org/10.3758/BF03213203

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage*, *63*(1), 240–244. https://doi.org/10.1016/j.neuroimage.2012.06.043

Martin, R., Ugen, S., & Fischbach, A. (Eds.). (2014). *Épreuves standardisées: Bildungsmonitoring für Luxemburg—Nationaler bericht 2011 bis 2013*. University of Luxembourg.

Martinez-Lincoln, A., Cortinas, C., & Wicha, N. Y. Y. (2015). Arithmetic memory networks established in childhood are changed by experience in adulthood. *Neuroscience Letters*, *0*, 325–330. https://doi.org/10.1016/j.neulet.2014.11.010

Martini, S. (2021). *The influence of language on mathematics in a multilingual educational setting*. University of Luxembourg.

Martini, S., Schiltz, C., Fischbach, A., & Ugen, S. (2021). Identifying math and reading difficulties of multilingual children: Effects of different cut-offs and reference groups. In *Identifying math and reading difficulties of multilingual children: Effects of different cut-offs and reference groups* (pp. 200–228). De Gruyter. https://doi.org/10.1515/9783110661941-011

McClain, L., & Huang, J. Y. S. (1982). Speed of simple arithmetic in bilinguals. *Memory & Cognition*, *10*(6), 591–596. https://doi.org/10.3758/BF03202441

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, *44*(1–2), 107–157. https://doi.org/10.1016/0010-0277(92)90052-J

McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition*, *4*(2), 171–196. https://doi.org/10.1016/0278-2626(85)90069-7

McClung, N. A., & Arya, D. J. (2018). Individual Differences in Fourth-Grade Math Achievement in Chinese and English. *Frontiers in Education*, *3*, 29. https://doi.org/10.3389/feduc.2018.00029

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Structural plasticity in the bilingual brain. *Nature*, *431*(7010), 757–757. https://doi.org/10.1038/431757a

Meck, W. H., Church, R. M., & Gibbon, J. (1985). Temporal integration in duration and number discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(4), 591–597. https://doi.org/10.1037/0097-7403.11.4.591

Meeuwissen, M., Roelofs, A., & Levelt, W. J. M. (2003). Planning levels in naming and reading complex numerals. *Memory & Cognition*, *31*(8), 1238–1248. https://doi.org/10.3758/BF03195807

Michael, E. B., & Gollan, T. H. (2005). Being and Becoming Bilingual: Individual Differences and Consequences for Language Production. In *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389–407). Oxford University Press.

Miller, K. F., Smith, C. M., Zhu, J., & Zhang, H. (1995). Preschool Origins of Cross-National Differences in Mathematical Competence: The Role of Number-Naming Systems. *Psychological Science*, *6*(1), 56–60.

Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development*, *2*(3), 279–305. https://doi.org/10.1016/S0885-2014(87)90091-8

Ministère de l'Éducation Nationale. (2022, April 20). Languages in Luxembourg Schools. https://men.public.lu/en/themes-transversaux/langues-ecole-luxembourgeoise.html

Miura, I. T., Kim, C. C., Chang, C.-M., & Okamoto, Y. (1988). Effects of Language Characteristics on Children's Cognitive Representation of Number: Cross-National Comparisons. *Child Development*, *59*(6), 1445. https://doi.org/10.2307/1130659

Miura, I. T., Okamoto, Y., Kim, C. C., Steere, M., & Fayol, M. (1993). First graders' cognitive representation of number and understanding of place value: Cross-national comparisons: France, Japan, Korea, Sweden, and the United States. *Journal of Educational Psychology*, *85*(1), 24–30. https://doi.org/10.1037/0022-0663.85.1.24

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*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyawaki, K., Jenkins, J. J., Strange, W., Liberman, A. M., Verbrugge, R., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, *18*(5), 331–340. https://doi.org/10.3758/BF03211209

Moeller, K., Klein, E., Fischer, M. H., Nuerk, H.-C., & Willmes, K. (2011). Representation of Multiplication Facts-Evidence for partial verbal coding. *Behavioral and Brain Functions*, *7*(1), 25. https://doi.org/10.1186/1744-9081-7-25

Moeller, K., Pixner, S., Zuber, J., Kaufmann, L., & Nuerk, H.-C. (2011). Early place-value understanding as a precursor for later arithmetic performance—A longitudinal study on numerical development. *Research in Developmental Disabilities*, *32*(5), 1837–1851. https://doi.org/10.1016/j.ridd.2011.03.012

Moeller, K., Shaki, S., Göbel, S. M., & Nuerk, H.-C. (2015). Language influences number processing – A quadrilingual study. *Cognition*, *136*, 150–155. https://doi.org/10.1016/j.cognition.2014.11.003

Moeller, K., Zuber, J., Olsen, N., Nuerk, H.-C., & Willmes, K. (2015). Intransparent German number words complicate transcoding – a translingual comparison with Japanese. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00740

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), Article 5109. https://doi.org/10.1038/2151519a0

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious Masked Priming Depends on Temporal Attention. *Psychological Science*, *13*(5), 416–424. https://doi.org/10.1111/1467-9280.00474

Naccache, L., & Dehaene, S. (2001). The Priming Method: Imaging Unconscious Repetition Priming Reveals an Abstract Representation of Number in the Parietal Lobes. *Cerebral Cortex*, *11*(10), 966–974. https://doi.org/10.1093/cercor/11.10.966

Negen, J., & Sarnecka, B. W. (2009). *Young children's number-word knowledge predicts their performance on a nonlinguistic number task*. https://escholarship.org/uc/item/1q03q75z

Negen, J., & Sarnecka, B. W. (2012). Number-Concept Acquisition and General Vocabulary Development. *Child Development*, *83*(6), 2019–2027. https://doi.org/10.1111/j.1467-8624.2012.01815.x

Nichols, E. S., Wild, C. J., Stojanoski, B., Battista, M. E., & Owen, A. M. (2020). Bilingualism Affords No General Cognitive Advantages: A Population Study of Executive Function in 11,000 People. *Psychological Science*, *31*(5), 548–567. https://doi.org/10.1177/0956797620903113

Notebaert, K., Pesenti, M., & Reynvoet, B. (2010). The neural origin of the priming distance effect: Distance-dependent recovery of parietal activation using symbolic magnitudes. *Human Brain Mapping*, *31*(5), 669–677. https://doi.org/10.1002/hbm.20896

Notebaert, W., Houtman, F., Opstal, F. V., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition*, *111*(2), 275–279. https://doi.org/10.1016/j.cognition.2009.02.002

Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition*, *82*(1), B25–B33. https://doi.org/10.1016/S0010-0277(01)00142-1

Nuerk, H.-C., Weger, U., & Willmes, K. (2002). A Unit-Decade Compatibility Effect in German Number Words. *Current Psychology Letters*, *2002/1, 7*. https://doi.org/10.4000/cpl.149

Nuerk, H.-C., Weger, U., & Willmes, K. (2004). On the Perceptual Generality of the Unit-Decade Compatibility Effect. *Experimental Psychology*, *51*(1), 72–79. https://doi.org/10.1027/1618-3169.51.1.72

Nuerk, H.-C., Weger, U., & Willmes, K. (2005). Language effects in magnitude comparison: Small, but not irrelevant. *Brain and Language*, *92*(3), 262–277. https://doi.org/10.1016/j.bandl.2004.06.107

Odic, D., Le Corre, M., & Halberda, J. (2015). Children's mappings between number words and the approximate number system. *Cognition*, *138*, 102–121. https://doi.org/10.1016/j.cognition.2015.01.008

Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, *28*(2), 191–230. https://doi.org/10.1017/S0142716407070117

Otheguy, R., García, O., & Reid, W. (2015). Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review*, *6*(3), 281–307. https://doi.org/10.1515/applirev-2015-0014

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., Majoubi, J., Balakrishnan, N., & Anders-Jefferson, R. T. (2024). Bilingualism, like other types of brain training, does not produce far transfer: It all fits together. *International Journal of Bilingualism*, 13670069231214599. https://doi.org/10.1177/13670069231214599

Parker Jones, ‘Ōiwi, Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., Lee, H. L., Ramsden, S., Gazarian, K., Prejawa, S., Seghier, M. L., & Price, C. J. (2012). Where, When and Why Brain Activation Differs for Bilinguals and Monolinguals during Picture Naming and Reading Aloud. *Cerebral Cortex*, *22*(4), 892–902. https://doi.org/10.1093/cercor/bhr161

Peng, P., Lin, X., Ünal, Z. E., Lee, K., Namkung, J., Chow, J., & Sales, A. (2020). Examining the mutual relations between language and mathematics: A meta-analysis. *Psychological Bulletin*, *146*(7), 595–634. https://doi.org/10.1037/bul0000231

Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, *15*(2), 202–206. https://doi.org/10.1016/j.conb.2005.03.007

Piantadosi, S. T., Jara-Ettinger, J., & Gibson, E. (2014). Children's learning of number words in an indigenous farming-foraging group. *Developmental Science*, *17*(4), 553–563. https://doi.org/10.1111/desc.12078

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503. https://doi.org/10.1126/science.1102085

Pinel, P., Dehaene, S., Rivière, D., & LeBihan, D. (2001). Modulation of Parietal Activation by Semantic Distance in a Number Comparison Task. *NeuroImage*, *14*(5), 1013–1026. https://doi.org/10.1006/nimg.2001.0913

Pitt, B., Gibson, E., & Piantadosi, S. T. (2022). *Exact number concepts are limited to the verbal count range*. *33*(3), 371–381. https://doi.org/10.1177/09567976211034502

Pixner, S., Zuber, J., Heřmanová, V., Kaufmann, L., Nuerk, H.-C., & Moeller, K. (2011). One language, two number-word systems and many problems: Numerical cognition in the Czech language. *Research in Developmental Disabilities*, *32*(6), 2683–2689. https://doi.org/10.1016/j.ridd.2011.06.004

Poncin, A., Rinsveld, A. V., & Schiltz, C. (2020). *Units first or tens first: How bilingualism affects two-digit number transcoding?* PsyArXiv. https://doi.org/10.31234/osf.io/sg7ea

Poncin, A., Van Rinsveld, A., & Schiltz, C. (2019). Units-first or tens-first: Does language matter when processing visually presented two-digit numbers? *Quarterly Journal of Experimental Psychology*, *73*(5), 726–738. https://doi.org/10.1177/1747021819892165

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, *22*(3), 415–422. https://doi.org/10.1016/j.acn.2007.01.015

Powell, S. R., Berry, K. A., & Tran, L. M. (2020). Performance differences on a measure of mathematics vocabulary for english learners and non-english learners with and without mathematics difficulty. *Reading & Writing Quarterly*, *36*(2), 124–141. https://doi.org/10.1080/10573569.2019.1677538

Powell, S. R., Urrutia, V. Y., Berry, K. A., & Barnes, M. A. (2022). The word-problem solving and explanations of students experiencing mathematics difficulty: A comparison based on dual-language status. *Learning Disability Quarterly*, *45*(1), 6–18. https://doi.org/10.1177/0731948720922198

Power, R. J. D., & Dal Martello, M. F. (1990). The dictation of Italian numerals. *Language and Cognitive Processes*, *5*(3), 237–254. https://doi.org/10.1080/01690969008402106

Prior, A., Katz, M., Mahajna, I., & Rubinsten, O. (2015a). Number word structure in first and second language influences arithmetic skills. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00266

Prior, A., Katz, M., Mahajna, I., & Rubinsten, O. (2015b). Number word structure in first and second language influences arithmetic skills. *Frontiers in Psychology*, *6*(MAR), 266. https://doi.org/10.3389/fpsyg.2015.00266

Proios, H. (2002). Number representation deficit: A bilingual case of failure to access written verbal numeral representations. *Neuropsychologia*, *40*(13), 2341–2349. https://doi.org/10.1016/S0028-3932(02)00085-4

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859–862. https://doi.org/10.3758/BF03192979

Purpura, D. J., Napoli, A. R., Wehrspann, E. A., & Gold, Z. S. (2017). Causal Connections Between Mathematical Language and Mathematical Knowledge: A Dialogic Reading Intervention. *Journal of Research on Educational Effectiveness*, *10*(1), 116–137. https://doi.org/10.1080/19345747.2016.1204639

Purpura, D. J., & Reid, E. E. (2016). Mathematics and language: Individual and group differences in mathematical language skills in young children. *Early Childhood Research Quarterly*, *36*, 259–268. https://doi.org/10.1016/j.ecresq.2015.12.020

R Core Team. (2013). *R: A language and environment for statistical computing*. http://www.R-project.org/

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, *72*(2), 191–201. https://doi.org/10.1016/S0010-0277(99)00048-7

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic priming in number naming. *The Quarterly Journal of Experimental Psychology Section A*, *55*(4), 1127–1139. https://doi.org/10.1080/02724980244000116

Reynvoet, B., De Smedt, B., & Van den Bussche, E. (2009). Children's representation of symbolic magnitude: The development of the priming distance effect. *Journal of Experimental Child Psychology*, *103*(4), 480–489. https://doi.org/10.1016/j.jecp.2009.01.007

Ritchie, S. J., & Bates, T. C. (2013). Enduring Links From Childhood Mathematics and Reading Achievement to Adult Socioeconomic Status. *Psychological Science*, *24*(7), 1301–1308. https://doi.org/10.1177/0956797612466268

Roberts, W. A., & Mitchell, S. (1994). Can a pigeon simultaneously process temporal and numerical information? *Journal of Experimental Psychology: Animal Behavior Processes*, *20*(1), 66–78. https://doi.org/10.1037/0097-7403.20.1.66

Rodic, M., Zhou, X., Tikhomirova, T., Wei, W., Malykh, S., Ismatulina, V., Sabirova, E., Davidova, Y., Tosto, M. G., Lemelin, J.-P., & Kovas, Y. (2015). Cross-cultural investigation into cognitive underpinnings of individual differences in early arithmetic. *Developmental Science*, *18*(1), 165–174. https://doi.org/10.1111/desc.12204

RStudio Team. (2020). *RStudio: Integrated development environment for r* [Manual]. http://www.rstudio.com/

RTL - Radio Télé. (n.d.). RTL. Retrieved April 6, 2024, from https://radio.rtl.lu

Rugani, R., Regolin, L., & Vallortigara, G. (2008). Discrimination of small numerosities in young chicks. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*(3), 388–399. https://doi.org/10.1037/0097-7403.34.3.388

Rumbaugh, D. M., & Savage-Rumbaugh, S. (1987). *Summation in the Chimpanzee (Pan troglodytes)*. 9. https://doi.org/10.1037/0097-7403.13.2.107

Saad, L. (2010). *Transcodage des nombres chez l'enfant: Approche développementale, inter-linguistique et différentielle* (NNT: 2010DIJOL005). Université de Burgogne.

Saalbach, H., Eckstein, D., Andri, N., Hobi, R., & Grabner, R. H. (2013). When language of instruction and language of application differ: Cognitive costs of bilingual mathematics learning. *Learning and Instruction*, *26*, 36–44. https://doi.org/10.1016/j.learninstruc.2013.01.002

Salillas, E., Barraza, P., & Carreiras, M. (2015). Oscillatory Brain Activity Reveals Linguistic Prints in the Quantity Code. *PLOS ONE*, *10*(4), e0121434. https://doi.org/10.1371/journal.pone.0121434

Salillas, E., & Carreiras, M. (2014). Core number representations are shaped by language. *Cortex*, *52*, 1–11. https://doi.org/10.1016/j.cortex.2013.12.009

Salillas, E., & Martínez, A. (2018). Linguistic Traces in Core Numerical Knowledge: An Approach From Bilingualism. In *Language and Culture in Mathematical Cognition* (pp. 173–196). Elsevier. https://linkinghub.elsevier.com/retrieve/pii/B9780128125748000080

Salillas, E., & Wicha, N. Y. Y. (2012). Early Learning Shapes the Memory Networks for Arithmetic: Evidence From Brain Potentials in Bilinguals. *Psychological Science*, *23*(7), 745–755. https://doi.org/10.1177/0956797612446347

Sarnecka, B., Kamenskaya, V., Yamana, Y., Ogura, T., & Yudovina, Y. (2007). From grammatical number to exact numbers: Early meanings of ‘one’, ‘two’, and ‘three’ in English, Russian, and Japanese. *Cognitive Psychology*, *55*(2), 136–168. https://doi.org/10.1016/j.cogpsych.2006.09.001

Sarnecka, B. W. (2017). *Early Number Knowledge In Dual-Language Learners From Low-SES Households*. https://escholarship.org/uc/item/0s87c79k

Sasanguie, D., Defever, E., Maertens, B., & Reynvoet, B. (2014). The Approximate Number System is not Predictive for Symbolic Number Processing in Kindergarteners. *Quarterly Journal of Experimental Psychology*, *67*(2), 271–280. https://doi.org/10.1080/17470218.2013.803581

Sasanguie, D., Defever, E., Van den Bussche, E., & Reynvoet, B. (2011). The reliability of and the relation between non-symbolic numerical distance effects in comparison, same-different judgments and priming. *Acta Psychologica*, *136*(1), 73–80. https://doi.org/10.1016/j.actpsy.2010.10.004

Sasanguie, D., & Reynvoet, B. (2014). Adults’ Arithmetic Builds on Fast and Automatic Processing of Arabic Digits: Evidence from an Audiovisual Matching Paradigm. *PLOS ONE*, *9*(2), e87739. https://doi.org/10.1371/journal.pone.0087739

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., & Smedt, B. D. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3), e12372. https://doi.org/10.1111/desc.12372

Schneider, R. M., Sullivan, J., Marušič, F., Žaucer, R., Biswas, P., Mišmaš, P., Plesničar, V., & Barner, D. (2020). Do children use language structure to discover the recursive rules of counting? *Cognitive Psychology*, *117*(June 2019), 101263. https://doi.org/10.1016/j.cogpsych.2019.101263

Secada, W. G. (1991). Degree of bilingualism and arithmetic program solving in hispanic first graders. *The Elementary School Journal*, *92*(2), 213–231. https://doi.org/10.1086/461689

Sella, F., Slusser, E., Odic, D., & Krajcsi, A. (2021). The emergence of children's natural number concepts: Current theoretical challenges. *Child Development Perspectives*, *15*(4), 265–273. https://doi.org/10.1111/cdep.12428

Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology*, *12*(3), 281–300. https://doi.org/10.1111/j.2044-835X.1994.tb00635.x

Shusterman, A., Slusser, E., Halberda, J., & Odic, D. (2016). Acquisition of the Cardinal Principle Coincides with Improvement in Approximate Number System Acuity in Preschoolers. *PLOS ONE*, *11*(4), e0153072. https://doi.org/10.1371/journal.pone.0153072

Siegler, R. S., & Mu, Y. (2008). Chinese Children Excel on Novel Mathematics Problems Even Before Elementary School. *Psychological Science*, *19*(8), 759–763. https://doi.org/10.1111/j.1467-9280.2008.02153.x

Siemann, J., & Petermann, F. (2018). Evaluation of the Triple Code Model of numerical processing—Reviewing past neuroimaging and clinical findings. *Research in Developmental Disabilities*, *72*, 106–117. https://doi.org/10.1016/j.ridd.2017.11.001

Singmann, H. (2021). *Mixed Model Reanalysis of RT data* [Computer software]. https://cran.r-project.org/web/packages/afex/vignettes/afex_mixed_example.html

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2020). *afex: Analysis of factorial experiments* [Computer software]. https://CRAN.R-project.org/package=afex

Sittig, O. (1921). Störung des Ziffernschreibens und Rechnens bei einem Hirnverletzten. *European Neurology*, *49*(5), 299–306. https://doi.org/10.1159/000190644

Spaepen, E., Coppola, M., Flaherty, M., Spelke, E., & Goldin-Meadow, S. (2013). Generating a lexicon without a language model: Do words for number count? *Journal of Memory and Language*, *69*(4), 496–505. https://doi.org/10.1016/j.jml.2013.05.004

Spaepen, E., Coppola, M., Spelke, E. S., Carey, S. E., & Goldin-Meadow, S. (2011). Number without a language model. *Proceedings of the National Academy of Sciences*, *108*(8), 3163–3168. https://doi.org/10.1073/pnas.1015975108

Spelke, E. S. (2017). Core Knowledge, Language, and Number. *Language Learning and Development*, *13*(2), 147–170. https://doi.org/10.1080/15475441.2016.1263572

Spelke, E. S., & Tsivkin, S. (2001a). Initial knowledge and conceptual change: Space and number. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (pp. 70–98). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.005

Spelke, E. S., & Tsivkin, S. (2001b). Language and number: A bilingual training study. *Cognition*, *78*(1), 45–88. https://doi.org/10.1016/S0010-0277(00)00108-6

Stancher, G., Rugani, R., Regolin, L., & Vallortigara, G. (2015). Numerical discrimination by frogs (Bombina orientalis). *Animal Cognition*, *18*(1), 219–229. https://doi.org/10.1007/s10071-014-0791-7

Starkey, P., & Cooper, R. G. (1980). Perception of Numbers by Human Infants. *Science*, *210*(4473), 1033–1035. https://doi.org/10.1126/science.7434014

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, *110*(45). https://doi.org/10.1073/pnas.1302751110

Stein, M., Federspiel, A., Koenig, T., Wirth, M., Strik, W., Wiest, R., Brandeis, D., & Dierks, T. (2012). Structural plasticity in the language system related to increased second language proficiency. *Cortex*, *48*(4), 458–465. https://doi.org/10.1016/j.cortex.2010.10.007

Steiner, A. F., Banfi, C., Finke, S., Kemény, F., Clayton, F. J., Göbel, S. M., & Landerl, K. (2021). Twenty-four or four-and-twenty: Language modulates cross-modal matching for multidigit numbers in children and adults. *Journal of Experimental Child Psychology*, *202*, 104970. https://doi.org/10.1016/j.jecp.2020.104970

Steiner, A. F., Finke, S., Clayton, F. J., Banfi, C., Kemény, F., Göbel, S. M., & Landerl, K. (2021). Language effects in early development of number writing and reading. *Journal of Numerical Cognition*, *7*(3), 368–387. https://doi.org/10.5964/jnc.6929

Stevenson, H. W., Lee, S.-Y., Chen, C., Stigler, J. W., Hsu, C.-C., Kitamura, S., & Hatano, G. (1990). Contexts of Achievement: A Study of American, Chinese, and Japanese Children. *Monographs of the Society for Research in Child Development*, *55*(1/2), i. https://doi.org/10.2307/1166090

Stocco, A., & Prat, C. S. (2014). Bilingualism trains specific brain circuits involved in flexible rule selection and application. *Brain and Language*, *137*, 50–61. https://doi.org/10.1016/j.bandl.2014.07.005

Sulpizio, S., Del Maschio, N., Fedeli, D., & Abutalebi, J. (2020). Bilingual language processing: A meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *108*, 834–853. https://doi.org/10.1016/j.neubiorev.2019.12.014

Swanson, H. L., Arizmendi, G. D., & Li, J.-T. (2022). What mediates the relationship between growth in math problem-solving and working memory in English language learners? *Journal of Educational Psychology*, *114*(7), 1608–1632. https://doi.org/10.1037/edu0000718

Swanson, H. L., Kong, J., & Petcu, S. (2018). Math difficulties and working memory growth in english language learner children: Does bilingual proficiency play a significant role? Language. *Speech & Hearing Services in Schools*, *49*(3), 379–394. https://doi.org/10.1044/2018_LSHSS-17-0098

The Luxembourgish Education System – An Overview. (n.d.).

Tomoschuk, B., Ferreira, V. S., & Gollan, T. H. (2019). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition*, *22*(3), 516–536. https://doi.org/10.1017/S1366728918000421

Ugen, S., Martin, R., Reichert, M., Lorphelin, D., & Fischbach, A. (2013). Einfluss des Sprachhintergrundes auf Schülerkompetenzen [Influence of Language Background on Students' Competencies. In *Ministère de l'Education nationale et de la Formation professionnelle, SCRIPT, & Université du Luxembourg, Unité de Recherche EMACS (Eds.), PISA 2012 – Nationaler Bericht Luxemburg [PISA 2012—National Report Luxembourg* (pp. 100–113). MENJE.

Ugen, S., Schiltz, C., Fischbach, A., & Pit-ten Cate, I. M. (2021). *Introduction: Les troubles des apprentissages dans un contexte multilingue – un vrai défi*. Melusina Press. https://www.melusinapress.lu/read/introduction-les-troubles-des-apprentissages-dans-un-contexte-multilingue-un-vrai-defi/section/a2d3cbda-6532-408f-b87d-34e34d1058dd#index.xml-body.1_div.1

Uller, C., Jaeger, R., Guidry, G., & Martin, C. (2003). Salamanders (Plethodon cinereus) go for more: Rudiments of number in an amphibian. *Animal Cognition*, *6*(2), 105–112. https://doi.org/10.1007/s10071-003-0167-x

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1–2), 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Van Assche, E., Duyck, W., & Hartsuiker, R. J. (2012). Bilingual Word Recognition in a Sentence Context. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00174

Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: A meta-analysis. *Psychological Bulletin*, *135*(3), 452–477. https://doi.org/10.1037/a0015329

van der Ven, S. H. G., Klaiber, J. D., & van der Maas, H. L. J. (2017). Four and twenty blackbirds: How transcoding ability mediates the relationship between visuospatial working memory and math in a language with inversion. *Educational Psychology*, *37*(4), 487–505. https://doi.org/10.1080/01443410.2016.1150421

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, *39*(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

van Marle, K., Chu, F. W., Li, Y., & Geary, D. C. (2014). Acuity of the approximate number system and preschoolers' quantitative development. *Developmental Science*, *17*(4), 492–505. https://doi.org/10.1111/desc.12143

Van Opstal, F., Gevers, W., De Moor, W., & Verguts, T. (2008). Dissecting the symbolic distance effect: Comparison and priming effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review*, *15*(2), 419–425. https://doi.org/10.3758/PBR.15.2.419

Van Rinsveld, A. (2015). *DO YOU SPEAK NUMBERS? THE RELATION BETWEEN LANGUAGE AND NUMERICAL COGNITION THROUGH THE PRISM OF BILINGUALISM AND CROSS-LINGUISTIC INVESTIGATIONS.*

Van Rinsveld, A., Brunner, M., Landerl, K., Schiltz, C., & Ugen, S. (2015). The relation between language and arithmetic in bilinguals: Insights from different stages of language acquisition. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00265

Van Rinsveld, A., Dricot, L., Guillaume, M., Rossion, B., & Schiltz, C. (2017). Mental arithmetic in the bilingual brain: Language matters. *Neuropsychologia*, *101*, 17–29. https://doi.org/10.1016/j.neuropsychologia.2017.05.009

Van Rinsveld, A., & Schiltz, C. (2016). Sixty-twelve = Seventy-two? A cross-linguistic comparison of children's number transcoding. *British Journal of Developmental Psychology*, *34*(3), 461–468. https://doi.org/10.1111/bjdp.12151

Van Rinsveld, A., Schiltz, C., Brunner, M., Landerl, K., & Ugen, S. (2016). Solving arithmetic problems in first and second language: Does the language context matter? *Learning and Instruction*, *42*, 72–82. https://doi.org/10.1016/j.learninstruc.2016.01.003

Van Rinsveld, A., Schiltz, C., Landerl, K., Brunner, M., & Ugen, S. (2016). Speaking two languages with different number naming systems: What implications for magnitude judgments in bilinguals at different stages of language acquisition? *Cognitive Processing*, *17*(3), 225–241. https://doi.org/10.1007/s10339-016-0762-9

Vander Beken, H., & Brysbaert, M. (2018). Studying texts in a second language: The importance of test type. *Bilingualism: Language and Cognition*, *21*(5), 1062–1074. https://doi.org/10.1017/S1366728917000189

Vanhove, J. (2013). The Critical Period Hypothesis in Second Language Acquisition: A Statistical Critique and a Reanalysis. *PLoS ONE*, *8*(7), e69172. https://doi.org/10.1371/journal.pone.0069172

vanMarle, K., Chu, F. W., Mou, Y., Seok, J. H., Rouder, J., & Geary, D. C. (2018). Attaching meaning to the number words: Contributions of the object tracking and approximate number systems. *Developmental Science*, *21*(1), 1–17. https://doi.org/10.1111/desc.12495

Venkatraman, V., Siong, S. C., Chee, M. W. L., & Ansari, D. (2006). Effect of Language Switching on Arithmetic: A Bilingual fMRI Study. *Journal of Cognitive Neuroscience*, *18*(1), 64–74. https://doi.org/10.1162/089892906775250030

Verreyt, N., Woumans, E., Vandelanotte, D., Szmalec, A., & Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition*, *19*(1), 181–190. https://doi.org/10.1017/S1366728914000352

Verschaffel, L., Greer, B., & Corte, E. (2000). *Making Sense of Word Problems*. Swets & Zeitlinger.

Volmer, E., Grabner, R. H., & Saalbach, H. (2018). Language switching costs in bilingual mathematics learning: Transfer effects and individual differences. *Zeitschrift Für Erziehungswissenschaft*, *21*(1), 71–96. https://doi.org/10.1007/s11618-017-0795-6

Vukovic, R. K., & Lesaux, N. K. (2013). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology*, *115*(2), 227–244. https://doi.org/10.1016/j.jecp.2013.02.002

Wagner, K., Kimura, K., Cheung, P., & Barner, D. (2015). Why is number word learning hard? Evidence from bilingual learners. *Cognitive Psychology*, *83*, 1–21. https://doi.org/10.1016/j.cogpsych.2015.08.006

Wang, Y., Lin, L., Kuhl, P., & Hirsch, J. (2007). Mathematical and Linguistic Processing Differs Between Native and Second Languages: An fMRI Study. *Brain Imaging and Behavior*, *1*(3–4), 68–82. https://doi.org/10.1007/s11682-007-9007-y

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's Past is Prologue: Relations Between Early Mathematics Knowledge and High School Achievement. *Educational Researcher (Washington, D.C. : 1972)*, *43*(7), 352–360. https://doi.org/10.3102/0013189X14553660

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, *8*(3), 231–256. https://doi.org/10.1162/jocn.1996.8.3.231

Wicha, N. Y., Dickson, D. S., & Martinez-Lincoln, A. (2018). Arithmetic in the Bilingual Brain. In *Language and Culture in Mathematical Cognition* (pp. 145–172). Elsevier. https://doi.org/10.1016/B978-0-12-812574-8.00007-9

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

World Bank Open Data. (n.d.). World Bank Open Data. Retrieved April 6, 2024, from https://data.worldbank.org

World Health Organization. (2019). *ICD-11 for Mortality and Morbidity Statistics*. https://icd.who.int/browse/2024-01/mms/en#308101648

Woumans, E., Surmont, J., Struys, E., & Duyck, W. (2016). The Longitudinal Effect of Bilingual Immersion Schooling on Cognitive Control and Intelligence*. *Language Learning*, *66*(S2), 76–91. https://doi.org/10.1111/lang.12171

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155–193. https://doi.org/10.1016/0010-0277(90)90003-3

Wynn, K. (1992a). Addition and subtraction by human infants. *Nature*, *358*(6389), 749–750. https://doi.org/10.1038/358749a0

Wynn, K. (1992b). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251. https://doi.org/10.1016/0010-0285(92)90008-P

Xenidou-Dervou, I., Gilmore, C., van der Schoot, M., & van Lieshout, E. C. D. M. (2015). The developmental onset of symbolic approximation: Beyond nonsymbolic representations, the language of numbers matters. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00487

Xenidou-Dervou, I., van Atteveldt, N., Surducan, I. M., Reynvoet, B., Rossi, S., & Gilmore, C. (2023). Multiple number-naming associations: How the inversion property affects adults' two-digit number processing. *Quarterly Journal of Experimental Psychology*, 17470218231181367. https://doi.org/10.1177/17470218231181367

Xu, C., Di Lonardo Burr, S., Skwarchuk, S.-L., Douglas, H., Lafay, A., Osana, H. P., Simms, V., Wylie, J., Maloney, E. A., & LeFevre, J.-A. (2022). Pathways to learning mathematics for students in French-immersion and English-instruction programs. *Journal of Educational Psychology*, *114*(6), 1321–1342. https://doi.org/10.1037/edu0000722

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Ziegler, J. C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical Words are Read Differently in Different Languages. *Psychological Science*, *12*(5), 379–384. https://doi.org/10.1111/1467-9280.00370

Zuber, J., Pixner, S., Moeller, K., & Nuerk, H.-C. (2009). On the language specificity of basic number processing: Transcoding in a language with inversion and its relation to working memory capacity. *Journal of Experimental Child Psychology*, *102*(1), 60–77. https://doi.org/10.1016/j.jecp.2008.04.003

15 Erratum

In the depiction of the (B)TCM and when referring to the (B)TCM, the "approximate/semantic" code should read as "analogue/semantic magnitude" code (see p. 11, 12, Figure 2 and p. 258, Figure 8).

END