Learning and Instruction 42 (2016) 72–82

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Solving arithmetic problems in first and second language: Does the language context matter?

Amandine Van Rinsveld [a,*], Christine Schiltz [a], Martin Brunner [b], Karin Landerl [c], Sonja Ugen [d]

[a] Education, Culture, Cognition and Society Research Unit (ECCS) – FLSHASE, University of Luxembourg, Campus Belval, Maison des Sciences Humaines, 11, Porte des Sciences, L-4366 Esch-sur-Alzette, Luxembourg
[b] Free University of Berlin and Berlin-Brandenburg Institute for School Quality, e.V. Otto-von-Simson-Str. 15, 14195 Berlin, Germany
[c] Department of Psychology – Developmental Psychology, University of Graz, Universitätsplatz 2/DG, 8010 Graz, Austria
[d] Luxembourg Center for Educational Testing (LUCET), University of Luxembourg, Campus Belval, Maison des Sciences Humaines, 11, Porte des Sciences, L-4366 Esch-sur-Alzette, Luxembourg

ARTICLE INFO

Article history:
Received 24 October 2014
Received in revised form
5 January 2016
Accepted 6 January 2016
Available online xxx

Keywords:
Language
Context
Bilingualism
Arithmetic
Mathematics

ABSTRACT

Learning mathematics in a second language is a challenge for many learners. The purpose of the study was to provide new insights into the role of the language context in mathematic learning and more particularly arithmetic problem solving. We investigated this question in a German–French bilingual educational setting in Luxembourg. Participants with increasing bilingual proficiency levels were invited to solve additions in both their first and second instruction languages: German and French. Arithmetic problems were presented in two different conditions: preceded by a semantic judgment or without additional language context. In the French session we observed that additions were systematically performed faster in the condition with an additional language context. In contrast no effect of the context was observed in the German session. In conclusion, providing a language context enhanced arithmetic performances in bilinguals' second instruction language. This finding entails implications for designing optimal mathematic learning environments in multilingual educational settings.

1. Introduction

1.1. Bilingual learning

Being taught in another language than the native language occurs very often in bilingual or multilingual populations (e.g., Extra & Gorter, 2001). Additionally, language immersion teaching programs, where for instance mathematics, sciences or geography courses are taught in the learner's second language, became more and more popular over the last years because such programs are perceived as an efficient and natural way to increase a second language's proficiency (Johnson & Swain, 1997). Nevertheless, it is well-known that non-native learners are more likely to have academic difficulties in general and more particularly with mathematical word problems, at least partially due to their lower command of the instruction language (Abedi & Lord, 2001; Gross, Hudson, & Price, 2009; Kempert, Saalbach, & Hardy, 2011; PISA report, 2012). Bernardo and Calleja (2005) showed that bilingual children's proficiency in the language of instruction is one of the most important factors influencing their mathematic proficiency.

In fact, growing up and/or being taught in more than one language raises the question whether and how instruction language(s) impact(s) the learned content and more generally the learners' academic achievement. Immersion programs mainly assume that the learned contents are sufficiently language-independent to be transferred to the learners' *mental* language. However, large-scale studies indicate that instruction language may significantly impact the learned contents (Cuevas, 1984; PISA report, 2012). Experimental evidence show that bilinguals' learned contents might strongly be associated to the language in

* Corresponding author.
E-mail addresses: amandine.vanrinsveld@uni.lu (A. Van Rinsveld), christine.schiltz@uni.lu (C. Schiltz), Martin.Brunner@isq-bb.de (M. Brunner), karin.landerl@uni-graz.at (K. Landerl), sonja.ugen@uni.lu (S. Ugen).

http://dx.doi.org/10.1016/j.learninstruc.2016.01.003
0959-4752/

which they were learned. Marian and Kaushanskaya (2007) reported that bilinguals responded differently to general knowledge questions according to the language in which these questions were asked. For example, to "name a statue of someone standing with a raised arm while looking into the distance", bilinguals responded more likely "the Statue of Liberty" when the question was in English while they responded more likely "the statue of Mao" when the same questions was asked in Chinese. Even closer to academic learning situations, Marian and Fausey (2006) presented Chemistry, Biology, and History contents to bilinguals either in Spanish or in English and then asked them questions about these contents in both languages. Participants responded faster when the language of the test matched the language of instruction than when they had to switch between languages. Together, these studies show how strong learned contents might be associated to the language of instruction.

Other authors trained bilingual participants to calculate in one language and then asked them to solve calculations in both of their languages. Spelke and Tsivkin (2001) showed that bilinguals could transfer calculations trained in one language to another for approximate calculations, but not for exact calculations. These results were replicated on trained multiplications and subtractions in German–French bilinguals (Saalbach, Eckstein, Andri, Hobi, & Grabner, 2013). Moreover, Venkatraman, Siong, Chee, and Ansari (2006) found similar advantages for the trained language in bilinguals performing exact additions in comparison to the untrained language and also corresponding brain activation networks related to the use of language in exact calculation. Taken together, these studies shed light on the importance of the instruction language in arithmetic and suggest that memorized calculations are represented verbally in a language-specific format, which seems to be difficult to transfer to another language. Similarly to the other aforementioned academic domains (Marian & Fausey, 2006), mathematical contents may thus also strongly be associated to the language of instruction. More generally, language is thought to play an important role in mathematical learning.

1.2. Learning mathematics with language

Even though all humans have a pre-verbal ability to approximate numerical quantities, learning to count with number words is mandatory to acquire exact numerical representations and master more complex mathematical skills, including arithmetic (Gordon, 2004; Xu & Spelke, 2000). Exact calculation is thus typically a mathematical skill acquired through instruction and which is thought to rely on language at different levels. During the exact calculation process, working memory resources are needed to execute solving procedures, keep in memory the intermediate solutions in memory and update the final solution (Ashcraft, 1995; Hitch, 1978). Particularly, the phonological loop of Baddeley's working memory model (1992) provides a support to verbally repeat the numbers in mental calculation (Fürst & Hitch, 2000; Logie, Gilhooly, & Wynn, 1994). Moreover, once simple calculations (i.e. usually calculations with operands <10) become overlearned with extensive practice, they can be retrieved directly from rote long-term memory (Ashcraft, 1992; McCloskey, 1992), where they are thought to be stored in a language-specific format (Campbell, 1994; Dehaene & Cohen, 1995). Because language plays such a key role in numerical learning, it is highly relevant to investigate arithmetic problem solving in bilinguals.

In line with the above-mentioned importance of language in arithmetic, the literature provides recurrent evidence that bilinguals calculate faster and more accurately in their first and better mastered language than in their second language, in which they are generally less proficient (Frenck-Mestre & Vaid, 1993; Geary, Cormier, Goggin, Estrada, & Lunn, 1993; Marsh & Maki, 1976; McClain & Huang, 1982). Neuro-imaging studies on Chinese-English bilinguals suggested that the verbal code of the first language is recruited to perform calculation and that extra language processing is needed in second language calculation (Lin, Imada, & Kuhl, 2011; Wang, Lin, Kuhl, & Hirsch, 2007). Although, one study showed that highly proficient bilinguals could achieve similar arithmetic performance levels in both of their languages (Campbell & Epp, 2004). Bilinguals' arithmetic performances in each language might also evolve across time with increasing language proficiency (Van Rinsveld, Brunner, Landerl, Schiltz, & Ugen, 2015). Besides bilinguals' language proficiency, the language of mathematic instruction also seems to play an important role. Indeed, it was observed that bilinguals calculated better in their non-native instruction language than in their native language, which was not used in the context of mathematic education (Bernardo, 2001; Salillas & Wicha, 2012). These findings have implications for bilingual or immersion education programs because they highlight the fact that some mathematical skills like arithmetic involve verbal processes and consequently the language in which they are taught may drastically influence performances.

1.3. The role of context in language selection

A key aspect of bilinguals' language competences is their ability to select and access the appropriate language for each situation (see Abutalebi, 2008 for a review about language control and L2 representations). According to the *Bilingual Interactive Activation Plus* or BIA + model (Dijkstra & Van Heuven, 2002), bilinguals access word meanings (i.e., lexical access) in a non-selective way via a shared lexicon for both languages at an early stage of the process (i.e., the words identification system). Later, the appropriate output language is then chosen by a decision sub-process. Hence, the linguistic context of a current task is thought to work as a decision threshold at this later stage. Similarly in language production, Kroll, Bobb, Misra, and Guo (2008) suggest that bilinguals activate both of their languages in parallel to produce words, but simultaneously, they use inhibitory processes to only produce words in the right language. These inhibitory processes are thought to occur with different efficiency degrees according to the relative levels of language proficiencies. Indeed, the less a language is mastered (e.g. second language), the more difficult it will be to use it without being interfered by the better mastered language (e.g. native language). Interferences in the other direction are possible too but to a lesser extend. Thus, it is more difficult to refrain the dominant language during the use of the non-dominant language than the other way around (Bialystok, 2009).

Additionally to language proficiency, the specific language context of each situation plays a fundamental role in the language selection process. Grosjean proposed a theoretical framework to account for the importance of language context in bilingualism: the *language mode hypothesis* (for a summary of his own studies see Grosjean, 2001). According to the language mode hypothesis, bilinguals' first and second language activation levels vary along a monolingual-bilingual continuum depending on their current so-called language mode. The language mode is determined by several factors such as the linguistic environment, the linguistic demand of a task, the nature and the topic of the interaction, etc. The effect of language context on bilinguals' performances has been broadly investigated in word recognition or lexical decision tasks by using inter-lingual homographs, i.e. words that have two different meanings according to the language in which they are used, for example *gift* means poison in German but present in English (Dijkstra, Van Jaarsveld, & Brinke, 1998).

Thus, primed lexical decision tasks with inter-lingual homographs have been used to highlight the general effect of language context on bilinguals' word recognition (Elston-Güttler, Gunter, & Kotz, 2005; Paulmann, Elston-Güttler, Gunter, & Kotz, 2006). In these studies German–English bilinguals had to decide whether a target item was an L2 word or not (i.e., in English). Each target item was primed by an L2 word. When the prime was an inter-lingual homograph with an L1 meaning that was related to the target word in L2 (e.g., *gift* as a prime for *poison*, because the word *gift* means poison in L1), the lexical decision on the target word was taken faster than with priming words that had no meaning in L1. The aforementioned authors also presented a film in L1 or in L2 before the primed lexical decision tasks that remained always in L2. The results showed that presenting a film in L1 before the primed lexical decision task further enhanced this influence of inter-lingual homographs on the lexical decision. In contrast, presenting a film in L2 reduced the priming effect by making the meaning of the inter-lingual homographs in L1 less salient. Moreover, recognition of words primed by related L1 inter-lingual homographs words was improved when priming words were presented in sentences in comparison to isolated words, providing further evidence for the importance of the context in bilinguals' language processes (Paulmann et al., 2006).

In conclusion, Elston-Güttler et al. (2005) proposed the concept of *zooming into L2*: L2 language contexts set by the film and/or the sentences modulated the influence of L1 on the interpretation of the inter-lingual homographs. Relatedly, other studies have shown that the semantic content of the sentences can modulate the recognition or the translation of words in these sentences (Schwartz & Kroll, 2006; Van Hell & De Groot, 2008). These studies provide evidence for a remaining L1 activation during the L2 tasks, supporting the idea of a parallel activation of both languages (BIA + model). They also demonstrate that the language context surrounding a task has a major impact on performances, since it seems to facilitate the language selection process or the activation of the right language.

1.4. The present study

While some studies reported neutral or beneficial effects of second language use for mathematics, in many other cases non-native instruction languages might strongly interfere with optimal learning and entail a cost for mathematic acquisition, namely arithmetic problem solving. Consequently it is critical to understand in detail how language influences arithmetic performances of bilingual learners at different levels of language proficiency and to find new ways to overcome these potential costs.

One aspect that has never been investigated so far is how bilinguals' performances in arithmetic might be improved by setting the problems in a language context. The priming effect of a language context on linguistic tasks has been extensively studied in bilinguals, although it has never been extended so far to any mathematical tasks. The aim of the present study was to fill this gap by investigating whether bilinguals of different language proficiency levels would benefit or not from a language context in arithmetic problem solving.

The present study was conducted in Luxembourg, which has implemented bilingual education at the national level. Languages of instruction are both German and French, as teaching is held exclusively in German during primary school while instruction language progressively switches to French during secondary school, so that the students become increasingly proficient both in German and French through their education. This allowed composing homogenous samples of bilingual participants in terms of exposure to German and to French at a certain age. In the current study, four samples of students from different grade levels of Luxembourgish secondary school and one sample of adult university students (who all had attended secondary school in Luxembourg) were carefully recruited in order to track increasing language proficiency.

Bilingual participants had to solve arithmetic problems in both German and French in order to track the performance differences across increasing bilingual proficiency levels. Addition problems were chosen because all arithmetic operations are not equal in terms of verbal processes: additions and multiplications especially rely on language (Lemer, Dehaene, Spelke, & Cohen, 2003). Moreover, we have seen that simple and complex arithmetic problems do not rely on the same solving strategies: direct retrieval of the solution vs. calculation, and verbal processes occurring in these different strategies might be influenced differentially by the language of the task. Addition is the easiest operation when considering both simple (i.e. operands <10) and complex (i.e. operands >10) problems, because even the complex addition problems requiring a calculation remain problems of small size compared to complex multiplication problems (e.g., $53 + 27$ leads to a smaller solution than 53×27). The arithmetic problems of the current study thus consisted in both simple and complex addition problems in order to look closer at the influence of language and language context effects on both types of solving procedures.

As the main purpose of the current study was to understand the impact of the language context on bilinguals' arithmetic at different bilingual proficiency levels, additions were presented in two separated language sessions. Moreover, in order to set the additions in a language context, we designed a task in which each addition was preceded by a sentence in the same language as the addition (*context condition*) and on which participants had to perform a semantic judgment. After the semantic judgment on the priming sentence, participants solved the addition in the same language (i.e., additions were presented visually but participants had to give their answers orally). We compared this context condition to a *no context* condition where participants had to solely solve additions in the instructed language without any context.

We expected better performances in German than in French for all participants, as German is acquired before French in the Luxembourgish school system and it also is the language in which arithmetic is learned (Bernardo, 2001; Salillas & Wicha, 2012). Nevertheless, this language-related difference in arithmetic problem solving should reduce with increasing bilingual proficiency across groups. In line with the effects of language context observed in bilinguals' linguistic performances, we hypothesized that the context condition should enhance the arithmetic performances in comparison to the no context condition in both languages. The language context created by the semantic judgment on the preceding sentences should help the bilinguals to activate the appropriate language to answer the additions. We expected the context to be helpful especially in French, as French was their less dominant language and given that the context should especially help in the less dominant language according to *the zooming into L2* concept of Elston-Güttler et al. (2005). These authors reported that word recognition in L2 was more efficient when the words were included in L2 sentences and when a film in L2 was presented before the task. The advantage of the context condition should be more prominent in complex than in simple additions because of the enhanced difficulty to solve complex arithmetic problems compared to simple arithmetic problems in bilinguals. Finally, we expected context advantage to decrease with increasing bilingual proficiency and thus increasing mastery of both instruction languages.

2. Method

2.1. Participants

A total of 193 bilingual participants were recruited for the present study. The sample was composed of 36 student of grade 7 (21 females; mean age of 12.2 years), 33 students of grade 8 (13 females; mean age of 13.2 years), 35 students of grade 10 (15 females; mean age of 15.5 years), 41 students of grade 11 (19 females; mean age of 16.4 years) and 48 adult university students (34 females; mean age of 22.4 years). All participants spoke Luxembourgish (an official language of Luxembourg which developed from a dialectal variant of German) or German as native language and attended Luxembourgish school system (in the higher academic track). Thus, all of them (including the adults) received primary school instruction in German and learned French as a foreign language, which implies that the language of early mathematical instruction was German. Amount of exposure to languages depended on the group: The younger students (grades 7 and 8) were relatively less exposed to French because they were only taught mathematics in French. The older students (grades 10 and 11) had been considerably more exposed to French, as this was their teaching language for mathematics and all of their other courses, except German and foreign language courses. Finally, the adult group had even more exposure to French as they had been attending the complete primary and secondary Luxembourgish school curriculum. The Luxembourgish curriculum requires reaching at least the level C1 from the *European Framework of Reference*[1] in French and in German at the end of secondary education, so that it can be assumed that the adults became increasingly proficient in both languages through their education. In terms of mathematic curriculum, all participants were assumed to have a full mastery of arithmetic operations at the time of the experiment, as arithmetic problem solving including additions are learned during primary education.[2]

2.2. Stimuli

Eighty-for two-operand addition problems were presented during the entire experiment. The set was composed of 28 one-digit simple additions ranging from 1 to 9 (e.g., 4 + 2) and 56 two-digit complex additions ranging from 12 to 98 (e.g., 56 + 32). We only included problems with a solution up to 99 and we excluded all additions including a zero or a repetition of the same digit within the operands or the solution and between both operands. The requirement of a carry to be solved (with or without carry) and the position of the largest operand (left or right) were counterbalanced for all additions. We additionally counterbalanced the problem size (small when the solution ranged between 30 and 60 or large when the solution ranged between 60 and 90) and distance between operands (small for distance < 40 or large for distance >40) for complex additions.

The stimulus set was split in four blocks that were allocated to both conditions: no context and with context; and to both task languages: German and French. Carry, position of the largest operand, problem size and distance were also counterbalanced within those four blocks of stimuli. Seven training items preceded the 21 additions of each block. In the condition with context, a semantic judgment task was performed on sentences in the corresponding task-language. Participants were instructed to judge whether the statements were true or false in the sense as possibly true. Sentences were constructed as similar as possible in both task-languages, for instance: "*Un lapin mange une carotte*" and "*Ein Hase frisst eine Möhre*" (which means *a rabbit eats a carrot*), see 7. Appendix. The number of words in each sentence ranged from four to seven words and the average number of letters was about 22 letters per sentence. The number of letters and words in each sentence was carefully controlled between languages.

2.3. Procedure

Participants were tested individually in a quiet room at school for the students or at the university for the adults. Native language(s), the number of years spent in Luxembourgish schools, gender and linguistic background (under the form of self-rating of language proficiency) were checked in a short questionnaire before starting the experiment in order to ensure that all participants also had similar exposures to languages in these respects. For the self-rating of language proficiency, participants had to rate their German and French proficiencies on a scale ranging from 1 to 6 (the six levels of the scale corresponded to "insufficient", "sufficient", "satisfactory", "good", "very good", and "perfect").[3]

We ran the experiment on an Apple 13′ Macbook using Psyscope X B57 (Cohen, MacWhinney, Flatt, & Provost, 1993) where response times were recorded with a voice key on the Iolab USB Button Box. As the voice key only recorded the response onset, the experimenter wrote the solutions down and pressed a key to start the next trial, which started after an ISI of 500 ms. Additions appeared on a white screen in black (Arial, 90) until participants responded. Participants had to respond orally by pronouncing the solution into the microphone in the language of the task.

In the context condition, a written sentence was presented on the screen before each addition on a white screen in black (Arial, 50). This sentence was in the same language as the task-language for the additions and participants had to make a semantic judgment on it by orally answering if the sentence was true or false. The sentence remained on the screen until participants gave their response. After participant's answer to the sentence, addition appeared after a delay of 500 ms.

They were instructed to respond as accurately and as fast as possible. Responses were recorded with the microphone. The testing was organized in two language sessions: participants performed both conditions in one language and than the same tasks in the other language. Order of condition and task-language was randomized between participants. Instructions and interaction with the experimenter remained in German or in French, according to the language session. Adults received 20€ for their participation at the experiment. Informed consent was obtained from all participants. The entire experiment lasted about 50 min.

[1] http://www.men.public.lu/catalogue-publications/systeme-educatif/langues-ecole-luxembourgeoise/reajustement-enseignement-langues/fr.pdf, p. 45.

[2] http://www.men.public.lu/catalogue-publications/systeme-educatif/langues-ecole-luxembourgeoise/reajustement-enseignement-langues/fr.pdf, p. 28.

[3] In participants of grade 8, 10 and 11 (N = 106), we had access to information about their academic performances in German and French, so that we correlated their school scores with the self-rated language proficiency levels respectively in German, r = 0.36, $p < .001$ and in French, r = 0.58, $p < .001$. Results thus showed a positive relation between self-rated language proficiencies and academic performances.

Table 1
Means with standard errors of RT (ms) and CR (%) for simple and complex additions presented in both task languages with and without context.

		German		French		Total German & French	
		M	(SE)	M	(SE)	M	(SE)
		Complex Additions					
RT	No Context	3681	(105)	4784	(151)	4232	(121)
	With Context	3552	(93)	4451	(133)	4001	(105)
	Total	3616	(96)	4617	(136)	4117	(111)
CR	No Context	89.3	(0.7)	83.6	(1.0)	86.4	(0.7)
	With Context	88.5	(0.9)	84.6	(0.9)	86.6	(0.7)
	Total	88.9	(0.6)	84.1	(0.8)	86.5	(0.6)
		Simple Additions					
RT	No Context	1399	(33)	1659	(51)	1529	(38)
	With Context	1440	(34)	1586	(39)	1513	(33)
	Total	1420	(31)	1622	(40)	1521	(33)
CR	No Context	96.7	(0.5)	95.7	(0.6)	96.2	(0.4)
	With Context	97.1	(0.5)	95.7	(0.6)	96.4	(0.4)
	Total	96.9	(0.4)	95.7	(0.4)	96.3	(0.3)

		Total complex & simple	
		M	(SE)
RT	No Context	2881	(76)
	With Context	2757	(66)
	Total	2819	(69)
CR	No Context	91.3	(0.4)
	With Context	91.5	(0.5)
	Total	91.4	(0.4)

3. Results

Response times (RT) and correct response rates (CR) were collected. Before the analyses, training items and response times of non-correct responses were excluded from the dataset, and we also excluded RT of trials where the recorded onset was not an answer to a calculation. Our trimming procedure for the RT excluded all trials below or above three standard deviations from the mean of each participant and from the group mean.

We ran ANCOVAs on the RT and the CR with $\text{Context}_2 \times \text{Task language}_2 \times \text{Difficulty}_2$ as within-subject factors and Group_5 as between-subject factor. Context referred to the presence or the absence of a semantic judgment sentence preceding the additions. Task language was German or French (for instructions, sentences in the context condition, and production of the answers to the additions in both conditions) and difficulty corresponded to simple (one-digit) and complex (two-digit) additions. The group factor designated the age-group of the participants that could be: 7th graders, 8th graders, 10th graders, 11th graders or adults. Participants' gender and their language proficiency difference were introduced as covariates in the model. Indeed, even though the groups were carefully selected to be homogeneous in terms of native language, ages of language acquisition and language exposure, we could not exclude individual differences in language proficiency. Therefore the difference of self-rated language proficiency between German and French was computed (see Fig. 3 in appendix 6.2) and this variable was used as a covariate in order to control for individuals' variability in terms of relative proficiency in both languages. A Greenhouse-Geisser correction was applied when sphericity of the data was not assumed.

Table 1 shows that, in general, additions were performed faster and more accurately[4] when the task language was German than when it was French (RT: F (1,186) = 51.154; $p < .001$; $\eta^2 = 0.224$ & CR: F (1,186) = 9.847; $p = .002$; $\eta^2 = 0.053$). As expected, simple additions were performed faster and more accurately than complex additions (RT: F (1,186) = 401.263; $p < .001$; $\eta^2 = 0.694$ & CR: F (1,186) = 146.367; $p < .001$; $\eta^2 = 0.453$). Further, participants responded generally faster when the additions were presented with preceding semantic judgment (F (1,186) = 8.877; $p = .003$; $\eta^2 = 0.048$), but context did not impact the CR (F (1,186) = 0.430; $p = .513$; $\eta^2 = 0.002$).

The Language × Difficulty interaction was also significant in both RT RT (F (1,186) = 41.322; $p < .001$; $\eta^2 = 0.189$) and CR (F (1,186) = 5.026; $p = .026$; $\eta^2 = 0.028$). Pairwise comparisons showed that the difference between languages was greater in complex (F (1,184) = 50.122; $p < .001$; $\eta^2 = 0.221$) than in simple additions (F (1,184) = 22.974; $p < .001$; $\eta^2 = 0.115$), whereas the difference of CR between languages was only significant for complex additions (F (1,186) = 10.320; $p = .002$; $\eta^2 = 0.055$) but not for simple additions where CR were similar in both languages (F (1,186) = 1.070; $p = .302$; $\eta^2 = 0.006$).

Critically, the Language × Context interaction was significant in RT (F (1,186) = 9.183; $p = .003$; $\eta^2 = 0.049$), see Figs. 1 and 2. To decompose this interaction we ran pairwise comparisons revealing that in French, the presence of semantic judgments before additions made participants respond faster than without semantic judgment context (F (1,186) = 12.736; $p < .001$; $\eta^2 = 0.067$). However, in German, the semantic judgments did not significantly impact the RT (F (1,186) = 0.110; $p = .741$; $\eta^2 = 0.001$).

Additionally, the Context × Difficulty interaction was significant only for RT (F (1,186) = 10.178; $p = .002$; $\eta^2 = 0.054$), showing that (independent of the above-mentioned language effect) complex additions were performed faster with a semantic judgment before than without (F (1,186) = 12.380; $p = .001$; $\eta^2 = 0.065$), while in simple additions, the context did not impact the RT (F (1,186) = 0.225; $p = .636$; $\eta^2 = 0.001$), see Table 1.

Finally, we also observed a significant group effect on RT (F (4,186) = 13.018; $p < .001$; $\eta^2 = 0.227$) showing that the RT decreased across age-groups. Moreover, the group factor interacted

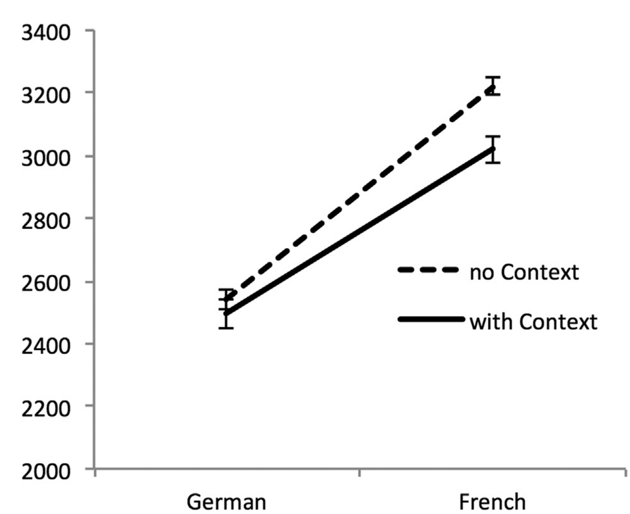

Fig. 1. Mean reaction times (ms) with standard error bars for the each task-language. Separated lines indicate the presence or not of a semantic judgment before each addition (context vs. no context conditions).

[4] As participants scored very high for some conditions, ceiling effects could have biased statistical analyses on CRs, see Table 2. The analyses were run again after that an arcsine transformation was applied to the data, showing the same results: difficulty effect, F (1,186) = 229.578; $p < .001$; $\eta^2 = 0.566$, language effect, F (1,186) = 12.473; $p = .001$; $\eta^2 = 0.066$, and interaction between language difficulty F (1,186) = 4.367; $p = .038$; $\eta^2 = 0.024$, while no other main effect of interaction reached significance (all Fs < 1 & ps > 0.05).

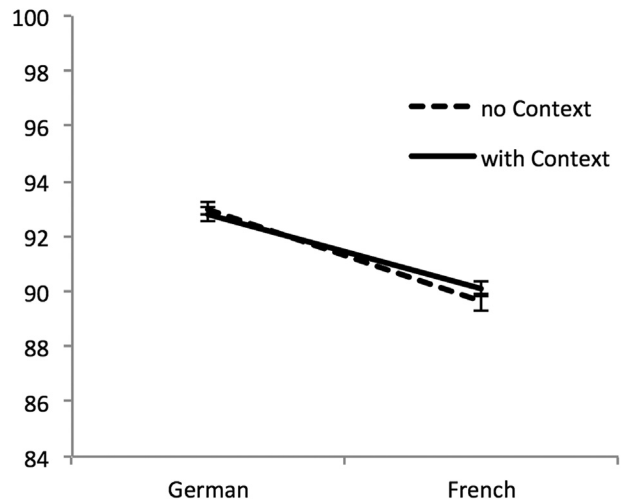

Fig. 2. Mean correct response rates (%) with standard error bars for the each task language. Separated lines indicate the presence or not of a semantic judgment before each addition (context vs. no context conditions).

with the task language, only for the RT (F (4,186) = 4.161; p = .003; η^2 = 0.086), showing that the response time differences between both task languages decreased across age-groups. See Table 2 for an illustration of these two last results. The only interaction implicating covariates we found was between the language proficiency difference and the difficulty × language interaction on RT (F (1,186) = 4.211; p = .042; η^2 = 0.023). The results from the pairwise comparisons showed that the language effect on RT was modulated by the covariate in complex (F (1,186) = 4.070; p = .045; η^2 = 0.022), but not in simple additions (F (1,186) = 0.555; p = .457; η^2 = 0.003). Concerning the CR, the covariate interacted with the language effect (F (1,186) = 5.482; p = .002; η^2 = 0.030). These interactions with the covariate thus showed that the individual differences in terms of language proficiency difference between languages observed within our participants could modulate the effects of the language on the performances. No other interactions reached significance and we observed no main effects of the covariates (all Fs < 1 & ps > 0.05).

In summary, as expected participants were faster and more accurate to solve simple compared to complex additions. They also needed less time and made fewer errors while calculating in German than in French and this language-difference was particularly pronounced for complex additions. While the effect of context also varied with task difficulty, the highly interesting context–language interaction indicated that participants especially benefited from the presence of a semantic task that preceded the addition problems during the French sessions.

4. Discussion

The present study focused on the role of language in bilinguals' arithmetic problem solving and addressed the question whether providing a language context could improve bilinguals' arithmetic performances in their initially and currently used languages of mathematic instruction. The study was conducted in Luxembourg, where instruction language switches from German to French in the school curriculum, so that students become increasingly proficient in both German and French over the years. This environment allowed the recruitment of five homogenous samples of participants in terms of exposure to German and to French at a certain age (7th, 8th, 10th and 11th graders and adults) who all acquired languages in a similar way, but were at distinct stages of this acquisition at the time of the study. Participants had to solve simple and complex additions that they had to answer orally in German or in French. To measure the potential effect of the language context, additions were presented in two different conditions: In the context condition, each addition was preceded by a sentence on which participants had to perform a semantic judgment. The sentence was presented in the same language as the subsequent addition. This condition was designed to activate German or French language before the addition was presented. This first condition was compared to a no context condition where additions were presented without any prior sentence.

4.1. Effects of the language on arithmetic problem solving

Results showed that all participants solved additions faster and more accurately in German than in French. This ***language effect*** occurred in all groups but the difference of response times between languages decreased with increasing bilingual proficiency, without disappearing completely even in the most proficient group (the adults). Thus, even though the participants of this study had become increasingly proficient in both French and German through their education, they still had a slight advantage in German for solving the additions. In the current bilingual population, German was both their dominant language and their language of instruction at the time of early mathematic acquisition. The current study thus replicates the findings that even highly proficient bilinguals remain more accurate and faster for

Table 2
Means with standard errors of RT (ms) and CR (%) for both task languages in each age-group of participants. Skewness and kurtosis of the distributions of RT and CR with standard error are provided for each age-group of participants.

		German		French		Total		Skewness	(SE)	Kurtosis	(SE)
		M	(SE)	M	(SE)	M	(SE)				
RT	7th graders	3222	(137)	4149	(193)	3686	(158)	0.405	(0.4)	0.756	(0.8)
	8th graders	2637	(146)	3357	(205)	2997	(168)	0.944	(0.4)	0.965	(0.8)
	10th graders	2525	(141)	3145	(198)	2835	(163)	0.873	(0.4)	0.628	(0.8)
	11th graders	2282	(129)	2693	(181)	2488	(148)	0.792	(0.4)	0.205	(0.7)
	Adults	1923	(119)	2255	(167)	2089	(137)	0.803	(0.4)	1.052	(0.7)
	Total	2518	(60)	3120	(85)	2819	(69)	1.034	(0.2)	0.774	(0.4)
CR	7th graders	92.1	(0.9)	88.2	(1.1)	90.1	(0.9)	−0.211	(0.4)	0.735	(0.8)
	8th graders	92.9	(1.0)	89.6	(1.1)	91.3	(0.9)	−0.64	(0.4)	1.435	(0.8)
	10th graders	92.7	(1.0)	89.0	(1.1)	90.8	(0.9)	−0.681	(0.4)	0.364	(0.8)
	11th graders	93.1	(0.9)	90.3	(1.0)	91.7	(0.8)	−0.888	(0.4)	0.482	(0.7)
	Adults	93.8	(0.8)	92.4	(0.9)	93.1	(0.7)	−0.893	(0.3)	0.421	(0.7)
	Total	92.9	(0.4)	89.9	(0.5)	91.4	(0.4)	−0.626	(0.2)	0.197	(0.4)

performing arithmetic tasks in their dominant language (Frenck-Mestre & Vaid, 1993; Geary et al., 1993; Marsh & Maki, 1976; McClain & Huang, 1982) and/or in the language in which they acquired early mathematical skills (Bernardo, 2001; Salillas & Wicha, 2012; Van Rinsveld et al., 2015).

The effect of language – additions solved faster and more accurately in German than in French – was visible in additions of both difficulty levels but it was more prominent in complex additions, which require more elaborate solving procedures than in simple additions that are thought to be solved by directly retrieving solutions from memory. These results consequently sustain the view that the solutions of simple problems might be stored in long-term memory under a verbal format (Campbell, 1994; Dehaene & Cohen, 1995) that is language-specific and makes the retrieval in another language than the language of encoding harder (Spelke & Tsivkin, 2001 see also Saalbach et al., 2013; Venkatraman et al., 2006). Further, the present results showed that not only retrieval of learned solutions but also knowledge about the solving procedures required by complex additions seems to be affected by the language of the task in bilinguals, suggesting that some steps of the complex addition solving rely on verbal processes that are difficult to transfer to another language than the language in which they were initially learned. Solving of complex problems might likely involve verbal processes of both types: retrieval of arithmetic facts combined with the application of learned solving procedures. Previous literature highlighted that working memory components are necessary to manage the successive steps of complex solving procedure and that these components might be at least partially of verbal nature (Ashcraft, 1995; Fürst & Hitch, 2000; Hitch, 1978; LeFevre et al., 2001).

Importantly, we observed these results even though we controlled for individual differences in terms of language proficiency difference between German and French. Although our groups of participants were homogenous samples in terms of school exposure to German and to French at a certain age, we asked them to rate their proficiency in each language and calculated the difference as a measure of their relative proficiency in these languages. Our results showed that this variable interacted with language effects on arithmetic performance in the sense that the larger their self-rated proficiency difference between German and French was, the larger were the language effects on RTs in complex additions and on CRs in both difficulty levels.

Taken together these results confirm the implication of language in arithmetic and highlight that arithmetic problems belong to the learned contents that are not transferrable from one language to another without any cognitive costs (Marian & Fausey, 2006; Spelke & Tsivkin, 2001 see also Saalbach et al. 2013; Venkatraman et al. 2006). Most importantly, the current results suggest that this limited transferability of arithmetic knowledge to another language in bilinguals concerns the retrieval of rote-memorized arithmetic contents (i.e., arithmetic facts) but also the implementation of procedural knowledge (i.e. solving processes).

4.1.1. Educational implications of language effects on arithmetic problem solving

In terms of implications for mathematical learning and instruction in bilingual education settings, these results remind us that even though attending bilingual education programs constitutes a wonderful opportunity to learn another language in a natural way, it also comes with costs because the learned contents are strongly associated to the language of instruction, especially in the case of mathematics. The current study provides evidence that when bilinguals have to perform very basic and supposedly fully mastered mathematical tasks (like addition problem solving) in their less dominant language, but which is also their current mathematic instruction language, their performances drop. This is especially the case for complex addition problems in comparison to simple addition problems and this effect seems to remain stable over the years of increasing bilingual proficiency. This closer look at bilinguals' basic arithmetic skills, which are the building blocks of more complex mathematical learning, highlights the importance of not considering mathematics and more generally number processing as “language-free” cognitive and scholastic tasks.

The academic difficulties for mathematical achievement recurrently reported in non-native learners have often been attributed unilaterally to the presence of word problems in mathematics, which are more difficult to understand in a less mastered language (Abedi & Lord, 2001; Gross et al., 2009; Kempert et al., 2011;). However, the current study draws the attention on an additional cause of these difficulties: the fact that basic mathematical skills, like arithmetic problems that do not contain any words, rely at least partially on verbal cognitive processing that might be more difficult to handle in the less mastered language. The difficulty of doing mathematics in a language that differs from the initial language of instruction seems to be due at least partially to the verbal aspects involved in mathematics at different levels: language-specific encoded knowledge (i.e., arithmetic facts) and language-specific procedural processing (i.e., solving with computation), which both make mathematics difficult to transfer from one language to another.

In order to ensure successful mathematical learning, bilingual education programs should take those verbal aspects of mathematics into account and provide new ways to overcome the involved cognitive costs of this limited transferability. Beyond the scope of bilingual education, a better understanding of those verbal aspects of mathematics is relevant for any education professional confronted to children with language difficulties, such as non-native learners (e.g., Cuevas, 1984) or children with language learning impairments (e.g., Landerl, Bevan, & Butterworth, 2004). Indeed many children who encounter language difficulties for different reasons also have increased risks of poor acquisition of mathematics.

4.2. Effects of the context on arithmetic problem solving

Furthermore, bilinguals performed additions faster in the context condition than in the no context condition, suggesting that bilinguals' addition problem solving might be enhanced when a non-numerical semantic judgment task on sentences activates the language of the additions. The ***context effect*** was modulated by task difficulty in the sense that only complex additions were performed faster in the context condition, whereas simple additions were solved equally fast in both conditions. Even more importantly, the language of the task modulated the context effect. Indeed, when the task was performed in French, participants performed the additions faster in the context condition than in the no context condition. However, when the participants performed the task in German, there were no differences anymore between both conditions. Thus, the language context as designed and implemented in the current study helped participants to solve additions faster in general, but this was especially the case when the task was performed in French. In other words, language context helped bilingual participants to solve additions faster in their less dominant language.

Interestingly, the interaction between context and language effect did not vary across the 5 age groups, even though increasing age groups systematically corresponded with higher bilingual proficiency levels. Moreover, participants' self-rated difference of language proficiency in both languages did not modulate the context effects. Thus, the aid provided by the language context when additions had to be solved in French was neither erased by the fact that the different age-groups had increased proficiency in both instruction languages over the years nor by the language proficiency differences among individuals. Taken together, the effect of the language observed throughout the age-groups and the modulation of arithmetic problem solving speed induced by the language context support the importance of language in learning and performing exact arithmetic.

The effects of the language context observed in the current study are in line with previous findings from the field of bilingual literature, which indicate that language selection processes come into play when the appropriate language has to be selected in bilinguals' language production (Abutalebi, 2008; Bialystok, 2009; Kroll et al., 2008). The language context that we created with the semantic judgments preceding the arithmetic task helped bilinguals to choose the appropriate language to solve additions faster. Our results thus fit Grosjean's *language mode hypothesis* (Grosjean, 2001) postulating that gradual activation levels of bilingual's first and second languages depend on the current language mode set by external factors such as the linguistic demand of the task and the language context. Indeed, the semantic judgments of the context condition seemed to increase the activation level of the language of the task, i.e. the language in which participants had to give their answer to the additions, so that participants were in a stronger *language mode* in the context condition than they were in the no context condition.

Moreover, as the context effect affected especially bilinguals' less dominant language (i.e., French), our findings extend the idea that a language context may help to *zoom into L2* (Elston-Güttler et al., 2005) to the field of arithmetical problem solving. Nevertheless, it is important to remember that the concept of *zooming into L2* by Elston-Güttler et al. (2005) intends to explain context effect on L2 word recognition among inter-lingual homograph tasks. Even though the language context of the task may not affect mathematical problem solving in exactly the same way as word recognition, we embrace the same idea that language context helps to perform a task especially in bilingual's less dominant language. To the best of our knowledge, the present study provided the first evidence of a positive effect of context on bilinguals' performances in a mathematical task.

When the additions were presented in French, which was participants' less dominant language, we cannot exclude that they mentally processed some of the verbal aspects of the task in their more dominant language and then translated the outcome to French before giving their answers. Indeed, literature examining the *inner speech* of bilingual individuals has highlighted that bilinguals' dominant language is crucial in verbal aspects of thinking. Although the use of a language for internal thoughts seems to depend on the proficiency in that language (e.g. De Guerrero, 1999), the results of Jiménez (2013) have shown that “bilinguals' dominant language played an important self-regulatory role in their verbalized thinking while the other language provided an extra set of cognitive resources and strategies that were employed when needed.” (p. 277). The use of another language than the dominant language for inner speech is thus possible in bilinguals under certain conditions, and we assume that the context condition of our experiment might have boosted the use of inner thinking in French. Namely for the case of complex additions presented with context, participants could have been more prone to process verbal working memory information directly in French, i.e. use French in the phonological loop for verbal rehearsal in arithmetic problem solving, because they were set in a more explicit “French mode” than in the condition without context where they just had to give the solutions to additions in French.

4.2.1. Educational implications of context effects

Crucially, our results indicate that presenting addition problems in a language context helps bilinguals to overcome the difficulty of doing the task in their less mastered language. These results imply that bilingual students may benefit from the presence of a constant language context when mathematics is taught in a non-dominant language or in a language that differs from early mathematic instruction language, because providing an adequate language context may enable them to better transfer acquired knowledge and procedures from one language to the other. In bilingual educational contexts like Luxembourg where both bilingual students and bilingual teachers are fluent in the same languages, language switching could occur very often in learning interactions (e.g., Moschkovich, 2006). Nevertheless, the language-context effects observed in the current study would rather suggest setting bilinguals in a unique *language mode* by providing a coherent language-context to enhance their performances in numerical tasks on which they will rely for building up more complex mathematics. Further research will be needed to verify this assumption by comparing bilingual learning environments where the language that is orally used in the classroom and written in handbooks is or is not kept constant throughout the teaching unit and by assessing whether one of both solutions is more helpful for students who are taught mathematics in a language that is not their dominant and/or their first language of mathematic acquisition.

Moreover, in terms of language learning, the challenge of performing complex tasks in another language than the dominant language ensures a progressive internalization of that language and by consequence an important gain in language proficiency (Centeno-Cortes & Jiménez, 2004; Jiménez, 2013). The ability to mentally process cognitive operations in a language, or in other words to *think in that language*, could be one of the steps of successful appropriation of a language (e.g. Lantolf & Yáñez, 2003). In the present case of mathematical teaching in a non-dominant language, it seems that to achieve successful learning of both the mathematical contents and the language of instruction, students should be encouraged to use the non-dominant language of instruction as a tool for thinking. Previous authors have already promoted teaching methods where the learned language becomes a mental support for thoughts (e.g. De Guerrero, 2004), namely in the case of problem solving activities (Jiménez, 2013). The current study extends these ideas to arithmetic problem solving and suggests that providing a corresponding language context, even with no direct link to the mathematical task, might enhance the recourse to the non-dominant language for numerical mental operations. Our results would suggest to the teachers to foster this aspect of language by overtly encouraging the learners to use the language of instruction as a tool for thought as much as possible.

4.3. Limitations

The suggestion to provide a language context should be considered with care because the academic language of mathematic may also constitute a hurdle for optimal mathematical learning in bilinguals' less mastered language. Research about the efficiency of test accommodation for bilingual learners showed that presenting mathematic problems with simplified linguistic instructions (Abedi & Lord, 2001) seemed to help them to overcome the linguistic complexity of mathematic problems. In the present study, the language context was intended to pre-activate the language of the addition problem solving task but it was not related to the task in terms of content (i.e. the sentences had nothing to do with mathematics). Further research will be needed to test the outcome of context effects on mathematics directly in the classroom, to understand how helpful the language context is in real-life mathematic classroom settings and to precise when bilinguals could benefit of a language context in mathematics (e.g. during initial learning stages or at later phases). One of the challenges will certainly be to find the right balance of providing a language context from which bilingual learners could benefit without adding linguistic complexity that may hamper mathematical achievement. Additionally, bilingual learners might benefit from an explicit training to transfer mathematical knowledge and procedures into the new instruction language before starting to learn more complex mathematics in this language. Appropriate training studies in bilingual school settings such as Luxembourg would be needed to validate this hypothesis.

4.4. Conclusions

The results of the current study support the view that language has a strong impact on arithmetic problem solving in bilinguals with different levels of proficiency. The verbal aspects implicated in mathematics may constitute a real hurdle that should be taken into consideration for bilingual education, but also more broadly for any learner with language weaknesses who may potentially benefit from the same kind of strategies that help overcoming those difficulties. Presenting arithmetic problems embedded in a matching language context seems to enhance performance by facilitating the transfer of knowledge to the new instruction language. Future research will allow to identify which are the most appropriate ways to implement helpful language contexts or which other help might allow learners to overcome language-related challenges during the complex real-life mathematical instruction.

Acknowledgment

The current research was funded by the Luxembourgish Fund for Scientific Research (FNR, Luxembourg) in the framework of the LangNum CORE project. The authors declare no conflict of interest that might be interpreted as influencing the research, and APA ethical standards were followed in the conduct of this work. We gratefully thank all the participants, the schools for their permission to recruit student participants and the student helpers for their support in the data collection. Correspondence and requests for materials should be addressed to Amandine Van Rinsveld (amandine.vanrinsveld@uni.lu).

Appendices

1. Sentence list of the semantic judgment task

French sentence	Letter count	German sentence	Letter count	Semantic judgment
“Un cheval est noir”	15	“Ein Pferd ist weiβ”	15	correct
“Un cheval vit dans l'étang”	21	“Ein Pferd lebt in dem Teich”	22	uncorrect
“Un juge travaille au restaurant”	27	“Ein Richter arbeitet im Restaurant”	29	uncorrect
“Un lion est un prédateur”	20	“Ein Löwe ist ein Raubtier”	21	correct
“Un lion est le roi des chèvres”	24	“Ein Löwe ist der König der Ziegen”	27	uncorrect
“Un vendeur est à la ferme”	20	“Ein Verkäufer ist auf dem Hof”	24	uncorrect
“Un lion rugit fort”	15	“Ein Löwe brüllt laut”	17	correct
“Un peintre repeint le mur”	21	“Ein Maler stricht die Wand”	22	correct
“Un cheval vit dans la ferme”	22	“Ein Pferd lebt auf dem Hof”	21	correct
“Un lapin mange une poule”	20	“Ein Hase frisst ein Huhn”	20	uncorrect
“Un juge porte un animal”	19	“Ein Richter trägt ein Tier”	22	uncorrect
“Un lion aboie fort”	15	“Ein Löwe bellt laut”	16	uncorrect
“Un lion est un rongeur”	18	“Ein Löwe ist ein Nagetier”	21	uncorrect
“Un renard est roux”	15	“Ein Fuchs ist rot”	14	correct
“Un fermier travaille à la ferme”	26	“Ein Bauer arbeitet auf dem Hof”	26	correct
“Un cheval est brun”	15	“Ein Pferd ist braun”	16	correct
“Un vendeur est à la caisse”	21	“Ein Verkäufer ist an der Kasse”	25	correct
“Un peintre utilise un pinceau”	25	“Ein Maler benutzt einen Pinsel”	26	correct
“Un coiffeur coiffe les cheveux”	26	“Ein Friseur frisiert die Haare”	26	correct
“Un musicien joue dans une caisse”	27	“Ein Musiker spielt in einer Kiste”	28	uncorrect
“Un peintre repeint le chat”	22	“Ein Maler stricht die Katze”	23	uncorrect
“Un vendeur travaille au magasin”	27	“Ein Verkäufer arbeitet im Laden”	27	correct
“Un cheval mange cette herbe”	23	“Ein Pferd frisst dieses Gras”	24	correct
“Un coiffeur coupe les animaux”	25	“Ein Friseur schneidet die Tiere”	27	uncorrect
“Un coiffeur coupe les cheveux”	25	“Ein Friseur schneidet die Haare”	27	correct
“Un juge travaille au tribunal”	25	“Ein Richter arbeitet am Gericht”	27	correct
“Un cheval est bleu”	15	“Ein Pferd ist blau”	15	uncorrect
“Un cheval est rouge”	16	“Ein Pferd ist gelb”	15	uncorrect
“Un dauphin est vert”	16	“Ein Delphin ist gelb”	16	uncorrect
“Un lion est le roi des animaux”	24	“Ein Löwe ist der König der Tiere”	26	correct
“Un architecte réalise un plan”	24	“Ein Architekt macht einen Plan”	26	correct
“Un architecte fait une maison”	25	“Ein Architekt baut ein Haus”	23	correct
“Un lapin mange cette herbe”	22	“Ein Hase frisst dieses Gras”	23	correct
“Un lapin vit dans le dessert”	23	“Ein Hase lebt in der Wüste”	21	uncorrect
“Un policier travaille dans la rue”	28	“Ein Polizist arbeitet in der Straβe”	30	correct
“Un policier porte un uniforme”	25	“Ein Polizist trägt eine Uniform”	27	correct
“Un policier travaille en classe”	27	“Ein Polizist arbeitet in der Klasse”	30	uncorrect
“Un renard vit dans la ferme”	22	“Ein Fuchs lebt auf dem Hof”	21	uncorrect
“Un dauphin mange un poisson”	23	“Ein Delphin frisst einen Fisch”	25	correct
“Un coiffeur coiffe les animaux”	26	“Ein Friseur frisiert die Tiere”	26	uncorrect
“Un architecte réalise un chien”	26	“Ein Architekt macht einen Hund”	26	uncorrect
“Un renard va dans son terrier”	24	“Ein Fuchs geht in seine Höhle”	24	correct

Note: The sentences were constructed with the same number of letters in each language. Sometimes for the color words, no literally translation was possible so that an equivalent color word was chosen instead, to respect the number of letters (e.g. “weiβ” means white but “noir” means black).

2. Self-rated language proficiency levels in German and in French

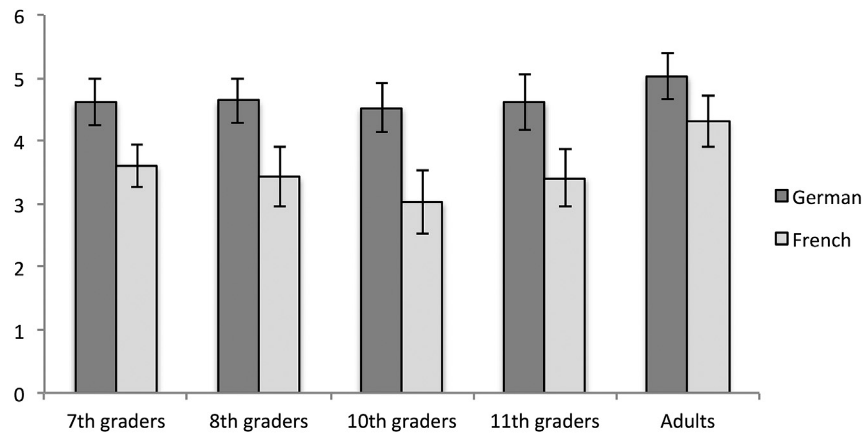

Fig. 3. Mean of the self-reported language proficiency ranked on a scale of 1–6 (corresponding to "insufficient", "sufficient", "satisfactory", "good", "very good", and "perfect") with standard error bars for German and French languages in each age-group.

References

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*(3), 219–234. http://dx.doi.org/10.1207/S15324818AME1403_2.

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

Ashcraft, M. H. (1992). Cognitive arithmetic: a review of data and theory. *Cognition, 44*(1–2), 75–106. http://doi.org/10.1016/0010-0277(92)90051-I.

Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: a review and summary of new directions. *Mathematical Cognition, 1*(1), 3–34.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556–559. http://dx.doi.org/10.1126/science.1736359.

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. http://dx.doi.org/10.3758/BF03195759.

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. http://doi.org/10.3200/GNTP.166.1.117-129.

Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(01), 3. http://dx.doi.org/10.1017/S1366728908003477.

Campbell, J. I. D. (1994). Architectures for numerical cognition. *Cognition, 53*(1), 1–44. http://doi.org/10.1016/0010-0277(94)90075-2.

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. http://dx.doi.org/10.1037/h0087447.

Centeno-Cortes, B., & Jimenez Jimenez, A. F. (2004). Problem-solving tasks in a foreign language: the importance of the L1 in private verbal thinking. *International Journal of Applied Linguistics, 14*(1), 7–35. http://doi.org/10.1111/j.1473-4192.2004.00052.x.

Cohen, J., MacWhinney, 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, & Computers, 25*(2), 257–271. http://dx.doi.org/10.3758/BF03204507.

Cuevas, G. J. (1984). Mathematics learning in English as a second language. *Journal for Research in Mathematics Education, 15*(2), 134. http://dx.doi.org/10.2307/748889.

De Guerrero, M. (1999). Inner speech as mental rehearsal: the case of advanced L2 learners. *Issues in Applied Linguistics, 10*(1), 27–55.

De Guerrero, M. (2004). Early stages of L2 inner speech development: what verbal reports suggest1. *International Journal of Applied Linguistics, 14*(1), 90–112.

Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical cognition, 1*(1), 83–120.

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*(03). http://dx.doi.org/10.1017/S1366728902003012.

Dijkstra, T., Van Jaarsveld, H., & Brinke, S. T. (1998). Interlingual homograph recognition: effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1*(1), 51–66. http://dx.doi.org/10.1017/S1366728998000121.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research, 25*(1), 57–70. http://dx.doi.org/10.1016/j.cogbrainres.2005.04.007.

Extra, G., & Gorter, D. (2001). In Multilingual Matters (Ed.), *The other languages of Europe: Demographic, sociolinguistic and educational perspectives* (p. 454). illustrée. Retrieved from http://books.google.com/books?hl=fr&lr=&id=hvmy_skUPNYC&pgis=1.

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

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory & Cognition, 28*(5), 774–782. http://dx.doi.org/10.3758/BF03198412.

Geary, D. C., Cormier, P., Goggin, J. P., Estrada, P., & Lunn, M. C. E. (1993). Mental arithmetic: a componential analysis of speed-of-processing across monolingual, weak bilingual, and strong bilingual adults. *International Journal of Psychology, 28*(2), 185–201. http://dx.doi.org/10.1080/00207599308247184.

Gordon, P. (2004). Numerical cognition without words: evidence from amazonia. *Science, 306*(5695), 496–499. http://dx.doi.org/10.1126/science.1094492.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Malden, MA: Blackwell Publishers Ltd.

Gross, J., Hudson, C., & Price, D. (2009). *The long term costs of numeracy difficulties*. London (UK): Every Child a Chance Trust and KPMG.

Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology, 10*(3), 302–323. http://doi.org/10.1016/0010-0285(78)90002-6.

Jiménez, A. F. J. (2013). Private speech during problem-solving activities in bilingual speakers. *International Journal of Bilingualism*. http://dx.doi.org/10.1177/1367006913509902, 1367006913509902.

Johnson, R. K., & Swain, M. (1997). *Immersion education: International perspectives*. Cambridge: Cambridge University Press.

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. http://dx.doi.org/10.1037/a0023619.

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

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: a study of 8–9-year-old students. *Cognition, 93*(2), 99–125.

Lantolf, J. P., & Yáñez, M.-C. (2003). *Talking yourself into Spanish: Intrapersonal communication and second language learning* (pp. 97–109). Hispania.

LeFevre, J.-A., Lei, Q., Smith-Chant, B. L., & Mullins, D. B. (2001). Multiplication by eye and by ear for Chinese-speaking and English-speaking adults. *Can. J. Exp. Psychol., 55*(4), 277–284. http://dx.doi.org/10.1037/h0087374.

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

Lin, J.-F. L., Imada, T., & Kuhl, P. K. (2011). Mental addition in bilinguals: an fMRI study of task-related and performance-related activation. *Cerebral Cortex, 22*(8), 1851–1861. http://dx.doi.org/10.1093/cercor/bhr263.

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition, 22*(4), 395–410. http://dx.doi.org/10.3758/BF03200866.

Marian, V., & Fausey, C. M. (2006). Language dependent memory in bilingual

learning. *Applied Cognitive Psychology, 20*(8), 1025e1047. http://dx.doi.org/10.1002/acp.1242.
Marian, V., & Kaushanskaya, M. (2007). Language context guides memory content. *Psychonomic Bulletin & Review, 14*(5), 925–933. http://dx.doi.org/10.3758/BF03194123.
Marsh, L. G., & Maki, R. H. (1976). Efficiency of arithmetic operations in bilinguals as a function of language. *Memory and Cognition, 4*, 459–464.
McClain, L., & Huang, J. Y. S. (1982). Speed of simple arithmetic in bilinguals. *Memory & Cognition, 10*(6), 591–596. http://dx.doi.org/10.3758/BF03202441.
McCloskey, M. (1992). Cognitive mechanisms in numerical processing: evidence from acquired dyscalculia. *Cognition, 44*(1–2), 107–157. http://doi.org/10.1016/0010-0277(92)90052-J.
Moschkovich, J. (2006). Using two languages when learning mathematics. *Educational Studies in Mathematics, 64*(2), 121–144. http://dx.doi.org/10.1007/s10649-005-9005-1.
Paulmann, S., Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2006). Is bilingual lexical access influenced by language context? *NeuroReport, 17*(7), 727–731. http://dx.doi.org/10.1097/01.wnr.0000214400.88845.fa.
PISA. (2012). *Results: What 15-year-olds know and what they can do with what they know*. OECD Publishing. http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf.
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. http://dx.doi.org/10.1016/j.learninstruc.2013.01.002.
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. http://dx.doi.org/10.1177/0956797612446347.
Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55*(2), 197–212. http://dx.doi.org/10.1016/j.jml.2006.03.004.
Spelke, E., & Tsivkin, S. (2001). Language and number: a bilingual training study. *Cognition, 78*(1), 45–88. http://dx.doi.org/10.1016/S0010-0277(00)00108-6.
Van Hell, J. G., & De Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica, 128*(3), 431–451. http://dx.doi.org/10.1016/j.actpsy.2008.03.010.
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*, 265. http://dx.doi.org/10.3389/fpsyg.2015.00265.
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. http://dx.doi.org/10.1162/089892906775250030.
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. http://dx.doi.org/10.1007/s11682-007-9007-y.
Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*(1), B1–B11. http://dx.doi.org/10.1016/S0010-0277(99)00066-9.