Inversion effects on mental arithmetic in English- and Polish-speaking adults

PLEASE CITE THE PUBLISHED VERSION

https://doi.org/10.1177/1747021819881983

PUBLISHER

SAGE Publications

VERSION

AM (Accepted Manuscript)

Inversion effects on mental arithmetic in English- and Polish-speaking adults

Carolin A. Lewis*

Leibniz-Institut für Wissensmedien, Tuebingen, Germany, Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, and Department of Psychiatry and Psychotherapy, University of Tuebingen,, Germany

Julia Bahnmueller*

Leibniz-Institut für Wissensmedien, Tuebingen, Germany, and Department of Psychology and LEAD Graduate School and Research Network, University of Tuebingen, Germany

Marta Wesierska

Department of Psychology, University of York, United Kingdom and Department of Human Communication Sciences, University of Sheffield, United Kingdom

Korbinian Moeller

Leibniz-Institut für Wissensmedien, Tuebingen, Germany, Department of Psychology and LEAD Graduate School and Research Network, University of Tuebingen, Germany

Silke M. Göbel

Department of Psychology, University of York, United Kingdom

Author Note

Carolin A. Lewis, Neuro-Cognitive Plasticity Laboratory, Leibniz-Institut für Wissensmedien, Tuebingen, now at Department of Psychiatry and Psychotherapy, University of Tuebingen (Calwerstr. 14, 72076 Tuebingen), Germany; Julia Bahnmueller, Neuro-Cognitive Plasticity Laboratory, Leibniz-Institut für Wissensmedien, Tuebingen (Schleichstraße 6, 72076 Tübingen), Germany, Department of Psychology, University of Tuebingen, Germany, and LEAD Graduate School and Research Network, University of Tuebingen, Germany; Marta Wesierska, Department of Psychology, University of York, United Kingdom, now at the Department of Human Communication Sciences, University of Sheffield (S102TN), United Kingdom; Korbinian Moeller, Neuro-cognitive plasticity Laboratory, Leibniz-Institut für

Wissensmedien, Tuebingen (Schleichstraße 6, 72076 Tübingen), Germany and Department of Psychology, University of Tuebingen, Germany, and LEAD Graduate School and Research Network, University of Tuebingen, Germany; Silke M. Göbel, Department of Psychology, University of York (YO10 5DD), United Kingdom. *: shared first authorship.

Correspondence concerning this article should be addressed to Julia Bahnmueller, Leibniz-Institut für Wissensmedien, Schleichstraße 6, 72076 Tuebingen, Germany. Email: j.bahnmueller@iwm-tuebingen.de, Phone: +49 7071 979-263, Fax: +49 7071 979-100.

Abstract

In some languages the order of tens and units in number words is inverted compared to symbolic digital notation (e.g., German 23 → "*dreiundzwanzig",* literally: "*three-and-twenty"*). In other languages only teen numbers are inverted (e.g., English 17 → "*seventeen*"; Polish 17 → "*siedemnaście*" literally "*seventeen*"). Previous studies focused on *between* group comparisons of inverted and non-inverted languages and showed that number word inversion impairs performance on basic numerical tasks and arithmetic. In two independent experiments, we investigated whether number word inversion affects addition performance *within* otherwise non-inverted languages (Exp. 1: English, Exp. 2: Polish). In particular, we focused on the influence of inverted (I; English: teen numbers ≥ 13, Polish: numbers 11 to 19) and non-inverted (N) summands with sums between 13 and 39. Accordingly, three categories of addition problems were created: *N+N*, *N+I*, and *I+I* with problem size matched across categories. Across both language groups, we observed that problems with results in the *twenty* and *thirty* number range were responded to faster when only non-inverted summands were part of the problems as opposed to problems with one or two inverted summands. In line with this, the cost of a carry procedure was largest for two inverted summands. Results support the notion that both language-specific and language-invariant aspects contribute to addition problem solving. In particular though, regarding language-specific aspects, results indicate that inverted number word formation of teens influences place-value processing of Arabic digits even in otherwise non-inverted languages.

239 words

Keywords: number word inversion, mental arithmetic, addition, carry effect, place-value processing

Inversion effects on mental arithmetic in English- and Polish-speaking adults

In daily life various number formats are used with number words and Arabic digits being the most frequent. One important step in numerical development is to map number symbols successfully and reliably to their respective number words (e.g., Chinese: 四 = sì (4), German: 4 = vier). In most languages, the mapping between single-digit number symbols and number words does not pose a major problem because one symbol corresponds directly to one single word. However, for numbers beyond the single-digit number range, the transparency of the mapping between symbolic notation and the corresponding number word/s varies considerably between languages. Most Western languages use the Arabic digit notation characterized by its place-value structure. In particular, in the Arabic number system, the value of a digit is determined by its position in the digit string (e.g., $44 = \{4\} \times 10^1 + \{4\} \times 10^0$). Many number word systems acknowledge this place-value structure by providing specific words for the decade position [e.g., 44 = *cuarenta y cuarto* (Spanish) = *fyrtiofyra* (Swedish) = *forty-four*] or by explicitly stating the power in multiplicative terms, e.g., at the hundreds position [444 = *cuatrocientos cuarenta y cuatro* (Spanish) = *fyrahundrafyrtiofyra* (Swedish) = *four hundred forty-four*]. These number word systems are quite transparent because they correspond to the Arabic place-value structure, others however are not. A common lack of transparency is the inversion of the order of number words for tens and units with respect to the Arabic place-value structure in some languages [e.g., 23 = *dreiundzwanzig (German)* = tlieta u għoxrin *(Maltese)* = literally: *three-and-twenty*)]. Importantly, evidence is accumulating that this lack of transparency for multi-digit numbers in languages with inverted number words has detrimental effects on basic numerical as well as arithmetic performance (e.g., Ganayim & Ibrahim, 2014; Göbel, Moeller, Pixner, Kaufmann, & Nuerk, 2014; Nuerk, Weger, & Willmes, 2005).

So far, however, inversion effects have primarily been investigated in cross-cultural designs, comparing performance of participants speaking a language with inverted number words to those speaking a language with non-inverted number words. Thereby, negative influences of inverted number words were shown in basic numerical tasks, for example, magnitude comparison (e.g., Moeller, Shaki, Göbel, & Nuerk, 2015; Nuerk et al., 2005) and number transcoding (e.g., writing down a number to dictation; Imbo, Bulcke, Brauwer, & Fias, 2014; Krinzinger et al., 2011). With regard to transcoding, for instance, significantly more place-coding errors were observed for languages with inverted number words [e.g., the German number word *vierundzwanzig* (literally: four and twenty) transcoded as 42]. In a study in Austrian first graders (German-speaking) Zuber, Pixner, Moeller, and Nuerk (2009) reported that up to 50% of children's transcoding errors were inversion-related.

Moreover, inversion effects on more complex tasks such as number line estimation (Helmreich et al., 2011) or mental arithmetic have also been demonstrated (e.g., Brysbaert, Fias, & Noël, 1998; Colomé, Laka, & Sebastian-Galles, 2010; Göbel et al., 2014; Lonnemann & Yan, 2015). For mental arithmetic, inversion-related influences on the *carry effect* in addition (Deschuyteneer, De Rammelaere, & Fias, 2005; Fürst & Hitch, 2000) were investigated specifically. The carry effect is a robust effect describing the observation that response times are usually longer and error rates higher for addition problems requiring a carry procedure (e.g., in 16 + 27 = 43 when adding the units 6 + 7 = 13 adds up to a sum larger than 9, in which case the decade digit of the unit sum has to be carried to the sum of the decade digits) than for non-carry problems. Importantly, because the effect is still present when overall magnitude is matched between carry and non-carry conditions, the effect indicates that multi-digit summands are not processed holistically (e.g., their overall magnitude is processed). Instead, the effect indicates that the single digits of a multi-digit number are processed componentially and comply with the place-value structure of Arabic number system (see, e.g., Nuerk, Moeller, & Willmes, 2015).

Investigating inversion-related influences on the carry effect in a symbolic addition task with Arabic digits, Göbel et al. (2014) found that the carry effect was more pronounced in German- (with an inverted number word system) than in Italian-speaking (with a non-inverted number word system) 2nd graders. The authors concluded that it takes more time to clearly identify and keep track of place-value positions during a carry problem when the number word structure of inverted number words provides inconsistent positional information. Importantly, these influences of inverted number words were observed in a non-verbal task using Arabic digits indicating that number word information is processed in a highly automated manner. In addition, the study by Lonnemann and Yan (2015) showed that inversion-related influences on the carry effect occur not only in children which are still acquiring (ir)regularities of a certain number word system but also in highly skilled adults. In particular, Lonnemann and Yan (2015) found a larger carry effect for German- than for Chinese-speaking participants. Because the Chinese number word system is perfectly transparent (e.g., 242 is spoken as “ èr bǎi sì shí èr” 2 100 4 10 2 literally two hundred four ten two), this result provides further evidence that number word inversion complicates place-value integration during carry operations in languages with inverted number words. Thus, while the carry effect per se indicates that place-value information is processed in multi-digit addition, the studies by Göbel and colleagues (2014) and Lonnemann and Yan (2015) corroborate the notion that the inversion of number words influences place-value processing in a very specific way.

An often neglected (or at least not explicitly discussed) fact is that teen numbers vary considerably with respect to the transparency of number words and digital notation. Crucially, this is even true for languages which otherwise present with a (quite) transparent (i.e., non-inverted) number word system beyond the teen number range (e.g., English, Italian, French). As is the case for two-digit numbers above the teen number range, there are certainly languages in which teen number words quite closely correspond to digital notation [e.g., Chinese: shí sān (ten three); Tamil: pathin moonduru (tenty three)]. However, other languages show a variety of peculiarities that potentially complicate numerical place-value processing and its development such as:

(i) the *regularity of construction* of number words [e.g., Italian: undici (one ten) but diciotto (ten eight), French: seize (~sixteen) but dix-sept (ten seven)],

(ii) *exceptional cases* where teen number words do not convey any explicit place-value structure because they might be, for instance, adopted from historical roots [e.g., eleven and twelve in English; treize, quatorze in French],

(iii) how the *decade term* is expressed [e.g., Spanish: no specific decade term, simply the word for ten (diez) (e.g., dieciseis (ten and six)); Tamil: explicit decade term with the same structure as other decades (i.e., pathin ~ tenty, cf. Dowker, Bala, & Lloyd, 2008); English: suffix –teen; Polish: suffix -naście (~ -teen)],

(iv) the *resemblance of teen and decade number words* (e.g. 13/30: English: thirteen/thirty; Italian: tredici/trenta; French: treize/trente),

(v) and most relevant for the present study the *order of terms reflecting unit and decade digits*. While for some number word systems (almost) all two-digit numbers are inverted (e.g., German, Dutch: all numbers larger than 12 except multiples of 10), the inversion of at least some teen-numbers is very common in otherwise consistent non-inverted number word systems [e.g., English: teen-numbers larger than 12 are inverted: “fourteen” (meaning four and ten) instead of tenfour; Polish: all teen numbers are inverted: jedenaście (oneteen), dwanaście (twoteen), trzynaście (threeteen)].

There is evidence suggesting that transparency of teen numbers seems to matter for children’s numerical development. Dowker and colleagues (2008) compared Tamil- and/or English-speaking children. Tamil number words are highly transparent with respect to all two-digit numbers including teens. In contrast, English teen number words (>13) are inverted whereas number words larger than 19 are not. Trying to decrease the influence of cultural factors other than language (all children were London, UK residents), the authors observed specific advantages in children speaking Tamil with respect to written calculation abilities mostly involving sums between 11 and 19. While results of this study suggest that there are specific transparency effects in the teen number word structure in children, the question remained of whether or not the transparency of teen number words influences later arithmetic performance and/or numerical processing in highly skilled adults.

The Present Study

While previous studies focused on *between* group comparisons of inverted and (mostly) non-inverted languages, the phenomenon of inverted teen-numbers allowed us to investigate whether number word inversion affects arithmetic performance *within* otherwise non-inverted languages in the current study. In particular, in Experiment 1 we investigated inversion-related influences within the English number word system for which teen numbers from 13 to 19 are inverted [e.g., 16 → “*sixteen”* (meaning *six-and-ten*)], while numbers larger than 20 are not inverted [e.g., 21 → “*twenty-one*” (meaning *twenty-and-one*)]. Moreover, to generalize results observed in English-speaking participants to another language group, we conducted the very same experiment in a sample of native speakers of Polish (Experiment 2). In Polish all teen numbers including eleven and twelve are inverted whereas numbers larger than 20 are not inverted. We used an addition task to investigate the carry effect as a specific indicator of place-value processing while controlling for problem size differences between (inverted) teen- and higher (non-inverted) two-digit numbers. Moreover, we chose a verification version of a symbolic addition task to minimize verbal in- or output requirements.

To investigate inversion-related influences, we considered three conditions with *teens*, *twens* or *thirties* as results of symbolic additions. These three conditions were included because of the natural specificities of English and Polish number words regarding the inversion of summands and/or results within the respective number ranges. In the *teens* condition, results were always inverted, whereas in the *twens* and *thirties* conditions, results were never inverted. Finally, the *thirties* condition provided the opportunity to compare two different kinds of carry problems with each other, which differed in how many inverted summands they contained (see Methods for a detailed description of all conditions). Each condition incorporated both carry and no-carry problems as well as inverted (I; English: numbers 13 to 19, Polish: numbers 11 to 19) and non-inverted numbers (N). In particular, carry and no-carry problems were differentiated according to three categories of addition problems: *N+N*, *N+I*, and *I+I*. Within each condition, problem size was matched between stimulus categories (see Table 1 for an overview of stimulus categories for both English and Polish). Based on previous results of *between* language comparisons, we hypothesized that number word inversion should influence place-value integration in mental arithmetic *within* one and the same language, especially in addition problems requiring a carry operation. More specifically, for both experiments and for all three conditions (*teens*, *twens*, and *thirties*), we expected (1) a significant carry effect (i.e., increased RT, whenever a carry operation is needed); (2) an effect of stimulus category (i.e., RT increase the more summands within a problem are inverted); (3) and differential carry effects depending on the number of inverted summands in the respective non-carry stimulus categories. In particular, due to added inversion-related processing costs, we hypothesized less pronounced carry effects the more inverted summands within the respective non-carry stimulus category. The latter would indicate that inversion specifically influences place-value processing of two-digit numbers.

Table 1. All stimulus categories per condition for Experiment 1 (English) and Experiment 2 (Polish).

	ENGLISH					**POLISH**				
		Stimulus category	**N**	**Inversion**	**Example**	**Inversion**	**N**	**Stimulus category**		
TEENS	**1**	**carry N+N**	8	N + N = I	7 + 9 = 16	N + N = I	8	**carry N+N**	**1**	**TEENS**
	2	no-carry N+I	4	N + I = I	2 + 14 = 16	N + I = I	8	no-carry N+**I**	2*	
	3	no-carry N+N	4	N + **N** = I	5 + 11 = 16	N + **I** = I				
TWENS	**4**	**carry N+I**	12	N + I = N	8 + 19 = 27	N + I = N	12	**carry N+I**	**4**	**TWENS**
	5	no-carry I+I	4	I + I = N	13 + 14 = 27	I + I = N	8	no-carry I+I	5*	
	6	no-carry N+I	4	**N** + I = N	12 + 15 = 27	**I** + I = N				
	7	no-carry N+N	4	N + N = N	6 + 21 = 27	N + N = N	4	no-carry N+N	7	
THIRTIES	**8**	**carry I+I**	6	I + I = N	18 + 19 = 37	I + I = N	6	**carry I+I**	**8**	**THIRTIES**
	9	**carry N+N**	6	N + N = N	9 + 28 = 37	N + N = N	6	**carry N+N**	**9**	
	10	no-carry N+I	4	N + I = N	23 + 14 = 37	N + I = N	8	no-carry N+I	10*	
	11	no-carry 2N+2N[1]	4	N + **N** = N	25 + 12 = 37	N + **I** = N				
	12	no-carry 1N+2N[1]	4	N + N = N	6 + 31 = 37	N + N = N	4	no-carry N+N	12	

Note. [1] 2N+2N = no-carry problems with only non-inverted summands, for which both summands were two-digit summands; 1N+2N = no-carry problems with only non-inverted summands, which contained a one- and a two-digit summand.

Methods

Participants

For Experiment 1, 24 native-English speakers were tested at the University of York, UK (M_{age} = 20.63 years, SD_{age} = 2.46, range 18-29, 9 male). For Experiment 2, 28 native-Polish speakers were tested at the University of Silesia, Poland (M_{age} = 21.11 years, SD_{age} = 0.88, range 20-23, 4 male). All participants were right-handed and had normal or corrected to normal vision. Written informed consent was obtained from all individual participants included in the study. Both studies were approved by the ethics committee of the Department of Psychology, University of York (UK).

Stimuli and Design for all Conditions

Stimuli were identical for Experiment 1 and 2. A set of 64 critical addition problems with a correct solution probe (range of summands: 1-34; problem size: 13-39) was created (e.g., 2 + 14 = 16). For the same 64 problems, two different types of filler items were created for which an incorrect solution probe was presented that i) had a distance of ± 1 (e.g., 2 + 14 = 15) or ii) ± 2 from the correct result (e.g., 2 + 14 = 18). The 64 stimuli were presented twice with the correct result and once per filler type. Additionally, all addition problems were presented twice, with the position of summands reversed, resulting in a total set of 256 x 2 = 512 addition problems.

The 64 addition problems were further split into 3 conditions based on the number range of the results: *teens* (range 13-19; e.g., 2 + 15 = 16), *twens* (range 20-29; e.g., 13 + 14 = 27), and *thirties* (range 30-39; e.g., 23 + 14 = 37). Moreover, the need for a carry operation (carry vs. no-carry) and inversion of summands was manipulated (none vs. one vs. both of the summands inverted, see Table 1). Note that numbers 11 and 12 were classified as non-inverted in Experiment 1 (English) and as inverted in Experiment 2 (Polish). This is because the respective number words in Polish are inverted [i.e., jedenaście (oneteen) and dwanaście (twoteen)] while in English (i.e., eleven and twelve) they are not inverted and lack any explicit place-value structure. Within each condition, problem size of each of the two summands as well as of results/incorrect solution probes was matched. Additionally, problem size of decades and units of summands and results was matched separately per condition and between carry and no-carry items. Moreover, position (left/right) of the smaller addend within the problem, and the parity of the summands and the correct result were matched. Finally, multiples of ten were not included in the stimulus set[1]. See https://osf.io/9jvgw/ for a list of all critical addition problems used.

Procedure

The same procedure was used for Experiment 1 and 2. The experiment was presented on an 18” screen driven at a resolution of 1024 x 768 pixels and a refresh rate of 60 Hz. Participants sat approximately 50 cm in front of the screen. After giving their written informed consent, participants were instructed to respond as quickly and as accurately as possible with the right index finger (i.e., press the right “Ctrl” key) when the presented solution was correct, and with the left index finger (i.e., press the left “Ctrl” key) when the solution was incorrect. Stimuli were presented centrally in white colour (font: *Courier New*, bold; font size: 24) against a black background until a response was given or the time limit of 8s was reached. A fixation cross preceded each trial and was presented for 500 ms in the middle of the screen. Trials were separated by an interstimulus interval of 500 ms.

[1] Ties were also not included with the exception of number 11 which was necessary to keep problem size matched between stimulus categories for each condition.

The experiment started with six practice trials. The experiment was presented in two blocks (A and B) of 256 trials each which only differed with respect to the order of summands (the order was reversed in block B). Trial order was randomized within each block separately for each participant. Half of the participants started with block A, the other half started with block B. After every 64 trials, participants were given the opportunity to take a short break.

Analysis

Analyses focused on addition problems with correct results. Only RTs of correctly solved trials were analysed. Errors were infrequent and not further analysed (error rate Experiment 1: *M* = 4.5%, *SD* = 2.8%, Experiment 2: *M* = 5.7%, *SD* = 4.6%). A trimming procedure eliminated RTs below or above 3 standard deviations of a participant's mean. Trimming resulted in average loss of 1.6% (*SD* = 0.6%) of data in Experiment 1 and 1.4% (*SD* = 0.8%) in Experiment 2. To control for differences in general processing speed, RTs were *z*-transformed prior to all analyses. For a better understanding, results are also reported in ms. The overall results pattern for both Experiment 1 and Experiment 2 is presented in Figure 1A and Figure 1B, respectively. For reasons of clarity, we report the results separately for Experiment 1 and 2 and for each of the three conditions (i.e., *teens, twens, and thirties*).

Experiment 1 - English

Results -Teens

The *teen* condition in Experiment 1 contained inverted numbers both as summands and results and consisted of three stimulus categories: 1) carry addition problems (*carry*), and two stimulus categories of no-carry addition problems. These two no-carry categories only differed in how many inverted summands they contained: 2) no-carry problems with one inverted summand (*no-carry N+I*) and 3) no-carry problems with only non-inverted summands (*no-carry N+N*). Note, that a stimulus category containing two inverted summands in the *teen* condition is not possible with English number words, as the results would be larger than 20.

A repeated-measures ANOVA with the factor stimulus category revealed significant differences in overall RT between stimulus categories (see Figure 1A and Table 2). Bonferroni-Holm-corrected pair-wise comparisons indicated significant differences between all three stimulus categories. Importantly, the significant comparisons of *carry* and *no-carry N+N* as well as *carry* and *no-carry N+I* reflected the expected carry effects. A significant difference between *no-carry N+N* and *no-carry N+I* indicated that no-carry additions with one inverted summand (*N+I*) were responded to faster than no-carry problems which did not involve an inverted summand (*N+N*).

To directly investigate the underlying mechanisms that led to the observed RT differences between stimulus categories, we evaluated possible differences in the carry effect due to the involvement of an inverted summand. Therefore, we computed the respective carry effects by subtracting RT of the two no-carry categories from the carry category: *N+N* [carry] - *N+I* [no carry] and *N+N* [carry] - *N+N* [no carry]. A paired samples *t*-test showed a significant difference between the respective carry effects indicating a larger carry effect when no-carry problems were involved as compared to no inverted summand (Table 2B).

Results - Twens

We included the *twen* condition, because it involved inverted numbers only as summands and not as results. In Experiment 1, this condition consisted of four stimulus categories, one category with addition problems 4) with a carry operation (*carry*), and three stimulus categories without a carry operation (see Table 1). Complementarily to the *teen* condition, the three no-carry categories of the *twen* condition reflected all possible combinations of inverted and non-inverted summands: 5) no-carry problems with two inverted summands (*no-carry I+I*), 6) no-carry problems with one inverted summand (*no-carry N+I*), and 7) no-carry problems with only non-inverted summands (*no-carry N+N*). Note that all results of the *twen* condition were not inverted, which means that observed differences in RT would be driven by the inversion property of summands only.

Analyses of the *twen* condition followed the same logic as for the *teen* condition. A repeated-measures ANOVA with the factor stimulus category revealed significant differences in RT between stimulus categories (see Figure 1A and Table 2A). Bonferroni-Holm-corrected pair-wise comparisons indicated the expected carry effects: RTs for carry problems were significantly longer than those to all no-carry categories. Moreover, there were significant differences between all three *no-carry* categories. This indicates that no-carry additions involving only non-inverted summands (*N+N*) were responded to faster than no-carry additions involving one or two inverted summands.

We tested again possible differences in the carry effect due to the involvement of inverted summands. Respective carry effects were computed by subtracting RT of the no-carry categories from the carry category: *N+I* [carry] - *I+I* [no carry], *N+I* [carry] - *N+I* [no carry], and *N+I* [carry] - *N+N* [no carry]. A repeated-measures ANOVA with stimulus category as factor revealed significant differences between the respective carry effects (Table 2 B). Bonferroni-Holm-corrected pairwise comparisons indicated that the carry effect was stronger when no-carry problems involved only non-inverted summands (*N+N*), compared to one (*N+I*, $p < .001$) or two inverted summands (*I+I*, $p < .001$) (Table 2). The carry effects

N+I [carry] - *I+I* [no carry] and *N+I* [carry] - *N+I* [no carry] also differed significantly (p = .030).

Results - Thirties

The *thirties* condition in Experiment 1 also involved inverted numbers only as summands and not as results. In addition to the *twen* condition, the *thirties* condition allowed us to compare two categories of carry problems with each other: 8) one containing only inverted summands (*carry I+I*) and 9) one containing only non-inverted summands (*carry N+N*). Again, the *thirties* condition also involved no-carry problems, which differed in how many inverted summands they contained: 10) no-carry problems with one inverted summand (*no-carry N+I*), 11) no-carry problems with only non-inverted summands, for which both summands were two-digit summands (*no-carry 2N+2N*), and 12) no-carry problems with only non-inverted summands, which contained a one- and a two-digit summand (*no-carry 1N+2N*).

The ANOVA with the within-participant factor stimulus category revealed significant differences in RT between stimulus categories. Bonferroni-Holm-corrected pair-wise comparisons indicated significant differences between stimulus categories (Figure 1A and Table 2A). Significant comparisons of carry and no-carry categories reflected the expected carry effects with an exception for the comparison of *carry N+N* and *no-carry N+I*, which was not significant. Additionally, we observed significant differences between *no-carry N+I* and *no-carry 2N+2N* , *no-carry N+I* and *no-carry 1N+2N*, as well as between *no-carry 2N+2N* and *no-carry 1N+2N*. This reflected that no-carry additions involving only non-inverted summands (*N+N*) were responded to faster than no-carry additions involving one inverted summand (*N+I*). Moreover, no-carry additions with a one-digit summand and a two-digit summand (*1N+2N*) were responded to faster than no-carry additions, which only contained two-digit summands (*2N+2N*). Finally, the two different carry categories in the *thirties* condition differed significantly as well, with longer RTs for *carry I+I* (M = 0.68), i.e. for a carry problem containing two inverted summands, than for *carry N+N* (M = 0.22), i.e. a carry problem only containing non-inverted summands.

As the *thirties* condition involved two carry categories, the respective carry effects were calculated separately. Two separate repeated measures ANOVAs with stimulus category as factor yielded significant differences between the respective carry effects, both for *carry I+I* and *carry N+N* (Table 2B). In line with the results from the *twen* condition, Bonferroni-Holm-corrected pairwise comparisons indicated that the carry effect for *carry I+I* was also stronger, when no-carry problems involved non-inverted summands only compared to problems with one inverted summand (p = .001 for *I+I* [carry] - *N+I* [no carry] vs. *I+I* [carry] – *2N+2N* [no-carry]; p <.001 for *I+I* [carry] - *N+I* [no carry] vs. *I+I* [carry] – *1N+2N* [no-carry]). Moreover, the carry effect was also more pronounced for no-carry problems with only non-inverted summands, which contained a one- and a two-digit summand (*I+I* [carry] – *1N+2N* [no-carry]) as compared to no-carry problems with only non-inverted summands, for which both summands were two-digit summands (*I+I* [carry] – *2N+2N* [no-carry], p < .001). Similarly, Bonferroni-Holm-corrected pairwise comparisons indicated that the carry effect for *carry N+N* was stronger, when no-carry problems involved non-inverted summands only compared to problems with one inverted summand (p = .001 for *N+N* [carry] - *N+I* [no carry] vs. *N+N* [carry] – *2N+2N* [no-carry]; p <.001 for *N+N* [carry] - *N+I* [no carry] vs. *N+N* [carry] – *1N+2N* [no-carry]. Additionally, we observed a more pronounced carry effect for no-carry problems with only non-inverted summands, which contained a one- and a two-digit summand (*N+N* [carry] – *1N+2N* [no-carry]), compared to no-carry problems with only non-inverted summands, for which both summands were two-digit summands (*N+N* [carry] – *2N+2N* [no-carry], $p < .001$; Table 2B).

Table 2. Overview of results of Experiment 1 (English).

A Difference between stimulus categories.

		Stimulus category	Statistic	*p*-value	η_p^2	Pairwise comparisons			*p*-value
TEENS	**1**	**carry N+N**	$F(2,46) = 111.7$	< .001	.829	**carry N+N**	vs	no-carry N+N	< .001
	2	no-carry N+I				**carry N+N**	vs	no-carry N+I	< .001
	3	no-carry N+N				no-carry N+N	vs	no-carry N+I	.012
TWENS	**4**	**carry N+I**	$F(3,69) = 73.55$	< .001	.762	**carry N+I**	vs	no-carry I+I	< .001
	5	no-carry I+I				**carry N+I**	vs	no-carry N+I	< .001
	6	no-carry N+I				**carry N+I**	vs	no-carry N+N	< .001
	7	no-carry N+N				no-carry I+I	vs	no-carry N+I	.030
						no-carry I+I	vs	no-carry N+N	< .001
						no-carry N+I	vs	no-carry N+N	< .001
THIRTIES	**8**	**carry I+I**	$F(4,92) = 75{,}98$	< .001	.768	**carry I+I**	vs	**carry N+N**	< .001
	9	**carry N+N**				**carry I+I**	vs	no-carry N+I	< .001
	10	no-carry N+I				**carry I+I**	vs	no-carry 2N+2N	< .001
	11	no-carry 2N+2N[1]				**carry I+I**	vs	no-carry 1N+2N	< .001
	12	no-carry 1N+2N[1]				**carry N+N**	vs	no-carry N+I	.128
						carry N+N	vs	no-carry 2N+2N	< .001
						carry N+N	vs	no-carry 1N+2N	< .001
						no-carry N+I	vs	no-carry 2N+2N	.003
						no-carry N+I	vs	no-carry 1N+2N	< .001
						no-carry 2N+2N	vs	no-carry 1N+2N	< .001

B (Differences in) Carry effects for all three conditions.

	Carry effect	Statistic	*p*-value	η_p^2	z-RT (*SD*)	ms-RT (*SD*)
TEENS	carry N + N – no-carry N + I	$t(1,23) = 2.74$	< .012		1.00 (0.40)	788 (423)
	carry N + N – no-carry N + N				0.88 (0.42)	677 (378)
TWENS	carry N + I – no-carry I + I	$F(2,46) = 39.5$	< .001	.632	0.53 (0.37)	394 (319)
	carry N + I – no-carry N + I				0.65 (0.39)	480 (378)
	carry N + I – no-carry N + N				0.99 (0.40)	752 (366)
THIRTIES	carry I + I – no-carry N + I	$F(2,46) = 48.7$	< .001	.679	0.59 (0.39)	435 (312)
	carry I + I – no-carry 2N + 2N[1]				0.82 (0.38)	613 (321)
	carry I + I – no-carry 1N + 2N[1]				1.32 (0.45)	1015 (489)
	carry N + N – no-carry N + I	$F(2,46) = 48.7$	< .001	.679	0.14 (0.43)	143 (377)
	carry N + N – no-carry 2N + 2N[1]				0.37 (0.37)	321 (382)
	carry N + N – 1no-carry N + 2N[1]				0.87 (0.33)	723 (500)

Note. [1] 2N+2N = no-carry problems with only non-inverted summands, for which both summands were two-digit summands; 1N+2N = no-carry problems with only non-inverted summands, which contained a one- and a two-digit summand.

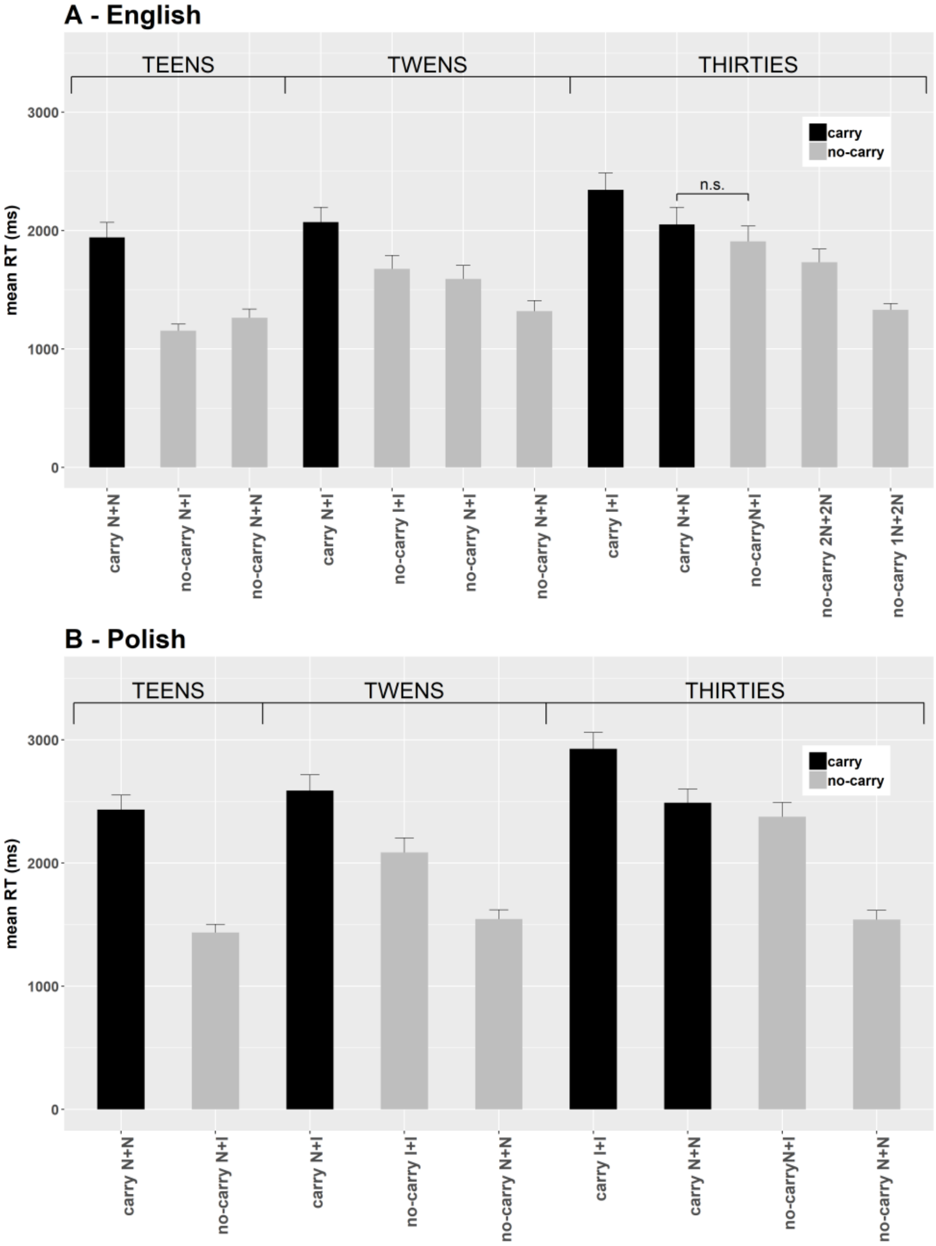

Figure 1. Mean RT of all stimulus categories per condition in ms (A) for Experiment 1 (English) and (B) for Experiment 2 (Polish). Error bars depict one standard error of the mean. For reasons of clarity, only the non-significant comparison is highlighted. All other comparisons between stimulus categories within each condition (*teens, twens, and thirties*) were significant.

Overall, results of Experiment 1 indicate that number word inversion at the item-level influences addition problem solving even within one and the same language group. However, for the teen number range, no-carry problems with one inverted summand (N+I) were solved faster than problems containing no inverted summand (N+N). This was opposite to our expectations. One possible explanation of this finding could be related to the stimuli we included in the *no-carry N+N* stimulus category for teens. This stimulus category used numbers 11 and 12 as one of the summands. Crucially and in contrast to all other teen number words, the number words eleven and twelve do not reference the units and tens explicitly.

Experiment 2 – Polish

Following up on the results of Experiment 1, the purpose of this second experiment was (i) to generalize the findings observed for the English-speaking sample to another language group and (ii) to replicate our findings in a language where all teen numbers are inverted. Thus, in Experiment 2 we tested a Polish-speaking sample because in Polish all teen numbers (including 11 and 12) are inverted. Note that because 11 and 12 are also inverted in Polish, we lose one stimulus category in each condition because stimulus categories 2 and 3, 5 and 6, as well as 10 and 11 do not differ in the number of inverted/non-inverted summands, respectively (see Table 1). Thus, for the teen number range we cannot investigate the influence of inverted teen numbers independently of the carry effect. For the twens and the thirties this s still possible.

Results -Teens

The *teen* condition in Experiment 2 contained inverted numbers both as summands and results and consisted of two stimulus categories: 1) carry addition problems (*carry*), and

2*) no-carry problems with one inverted summand (*no-carry N+I*). Note, that categories 2 and 3 of Experiment 1 merged together into one category in Experiment 2 (here category 2*), because 11 and 12 are inverted in Polish. We report analyses of these two stimulus categories first. Subsequently, for reasons of comparability with Experiment 1, we also compared *no-carry N+I* items including 11 and 12 to *no-carry N+I* items without 11 and 12 in the Polish data set. Moreover, as in English, a stimulus category containing two inverted summands in the *teen* condition is not possible with Polish number words either, as the results would be larger than 20.

For teen numbers, we observed a carry effect by directly comparing RTs of the stimulus categories *carry N+N* and *no-carry N+I* (see Figure 1B and Table 3A). As expected, no-carry additions were responded to faster than carry additions.

The subsequent comparison of *no-carry N+I* items including 11 and 12 to *no-carry N+I* items without 11 and 12 showed a significant difference, $t(1,27) = -6.45$, $p < .001$, indicating that – similar to Experiment 1 – no-carry additions with one inverted summand were responded to faster than no-carry problems with 11 and 12 as inverted summand.

Results – Twens

The *twen* condition in Experiment 2 consisted of three stimulus categories (see Table 1), one category with addition problems 4) with a carry operation (*carry*), and two stimulus categories without a carry operation: 5*) no-carry problems with two inverted summands (*no-carry I+I*), and 7) no-carry problems with only non-inverted summands (*no-carry N+N*). Note, that categories 5 and 6 of Experiment 1 merged together into one category in Experiment 2 (here category 5*), because 11 and 12 are inverted in Polish. Again, to report fully comparable analyses to Experiment 1, we additionally compared *no-carry I+I* items including 11 and 12 to *no-carry I+I* items without 11 and 12 in the Polish data set. As in English, all results of the *twen* condition were not inverted in Polish, which means that observed performance differences should be driven primarily by the inversion property of summands.

As in Experiment 1, a repeated-measures ANOVA with the factor stimulus category (3 categories) showed significant differences in RT between stimulus categories (see Figure 1B and Table 3A). The expected carry effects were confirmed by Bonferroni-Holm-corrected pair-wise comparisons: RTs for carry problems were significantly longer than those to both no-carry categories. The two no-carry categories also differed significantly, indicating that no-carry additions involving only non-inverted summands were responded to faster (*N+N*) than no-carry additions involving only inverted summands (*I+I*).

The respective carry effects were again computed by subtracting RT of the no-carry categories from the carry category: *N+I* [carry] - *I+I* [no carry], and *N+I* [carry] - *N+N* [no carry]. A paired *t*-test revealed a significant difference between the respective carry effects: the carry effect was stronger when no-carry problems involved only non-inverted summands (*N+N*), compared to two inverted summands (*I+I*) (Table 3B).

Moreover, we compared *no-carry I+I* items including 11 and 12 to *no-carry I+I* items without 11 and 12 and found, as expected, no significant difference, $t(1,27) = 0.51$, $p = .61$.

Results - Thirties

The *thirties* condition in Experiment 2 also involved inverted numbers only as summands and not as results (see Table 1). Comparable to Experiment 1, the *thirties* condition allowed us to compare two categories of carry problems with each other: 8) one containing only inverted summands (*carry I+I*) and 9) one containing only non-inverted summands (*carry N+N*). Again, the *thirties* condition also involved no-carry problems, which differed in how many inverted summands they contained: 10*) no-carry problems with one inverted summand (*no-carry N+I*), and 12) no-carry problems with only non-inverted summands, which contained a one- and a two-digit summand (*no-carry 1N+2N*). Note, that categories 10 and 11 of Experiment 1 form one category in Experiment 2 (here category 10*), because 11 and 12 are inverted in Polish. Additionally, we again compared *no-carry N+I* items including 11 and 12 to *no-carry N+I* items without 11 and 12.

The ANOVA with the within-participant factor stimulus category (3 categories) revealed the same pattern as already observed in Experiment 1. Again, a main effect of stimulus category was found. Bonferroni-Holm-corrected pair-wise comparisons indicated significant differences between stimulus categories (see Figure 1B and Table 3A). Significant comparisons of carry and no-carry categories reflected the expected carry effects. Additionally, we observed significant differences between *no-carry N+I* and *no-carry 1N+2N*. This reflected that no-carry additions with a one-digit and a two-digit non-inverted summand (*1N+2N*) were responded to faster than no-carry additions involving one inverted summand (*N+I*). Finally, the two different carry categories in the *thirties* condition differed significantly as well, with longer RTs for *carry I+I* , this means for a carry problem containing two inverted summands, than for *carry N+N* , this means a carry problem only containing non-inverted summands.

Two separate paired *t*-tests yielded significant differences between the respective carry effects, both for *carry I+I* and *carry N+N*. In line with the results from the *twen* condition, comparisons indicated that the carry effect for *carry I+I* was more pronounced for no-carry problems with only non-inverted summands, which contained a one- and a two-digit summand (*1N+2N*) compared to problems with one inverted summand (*N+I*) (*I+I* [carry] -

N+I [no carry] vs. *I+I* [carry] – *1N+2N* [no-carry]). Similarly, we observed a more pronounced *carry N+N* carry effect for no-carry problems with only non-inverted summands, which contained a one- and a two-digit summand (*1N+2N*) in contrast to no-carry problems with one inverted summand (*N+I*) (*N+N* [carry] - *N+I* [no carry] vs. *N+N* [carry] – *1N+2N* [no-carry].

Finally, we compared *no-carry N+I* items including 11 and 12 to *no-carry N+I* items without 11 and 12 and found the opposite pattern as in the teen condition: the significant difference between *no-carry N+I* items including 11 and 12 and *no-carry N+I* items without 11 and 12, $t(1,27) = 3.58$, $p = .001$, indicated that no-carry additions with 11 and 12 as inverted summand were responded to faster than no-carry problems with one inverted summand.

Table 3. Overview of results of Experiment 2 (Polish).

A Difference between stimulus categories.

	Stimulus category	Statistic	*p*-value	η_p^2	Pairwise comparisons			*p*-value
TEENS	**1 carry N+N**	t(1,27) = 16.22	< .001					
	2* no-carry N+I							
TWENS	**4 carry N+I**	F(2,54) = 106.9	< .001	.798	**carry N+I**	vs	no-carry I+I	< .001
	5* no-carry I+I				**carry N+I**	vs	no-carry N+N	< .001
	7 no-carry N+N				no-carry I+I	vs	no-carry N+N	< .001
THIRTIES	**8 carry I+I**	F(3,81) = 179.8	< .001	.869	**carry I+I**	vs	**carry N+N**	< .001
	9 carry N+N				**carry I+I**	vs	no-carry N+I	< .001
	10* no-carry N+I				**carry I+I**	vs	no-carry N+N	< .001
	12 no-carry N+N				**carry N+N**	vs	no-carry N+I	.047
					carry N+N	vs	no-carry N+N	< .001
					no-carry N+I	vs	no-carry N+N	< .001

B (Differences in) Carry effects for all three conditions.

	Carry effect	Statistic	*p*-value	z-RT (*SD*)	ms-RT (*SD*)
TEENS	carry N + N – no-carry N + I			0.90 (0.29)	998 (444)
TWENS	carry N + I – no-carry I + I	t(1,27) = -8.86	< .001	0.46 (0.36)	501 (426)
	carry N + I – no-carry N + N			0.94 (0.37)	1043 (511)
THIRTIES	carry I + I – no-carry N + I	t(1,27) = -15.41	< .001	0.51 (0.27)	552 (306)
	carry I + I – no-carry N + N			1.26 (0.36)	1386 (526)
	carry N + N – no-carry N + I	t(1,27) = -15.41	< .001	0.11 (0.28)	114 (309)
	carry N + N – no-carry N + N			0.87 (0.30)	948 (425)

Discussion

Specificities of number word formation (e.g., the inversion of number words with respect to Arabic digits) influence number processing in different tasks and paradigms, even when only Arabic digits are presented (e.g., number comparison: Nuerk et al., 2005; transcoding: Imbo et al., 2014; arithmetic: Göbel et al., 2014). Based on these previous results of *between* language comparisons, we hypothesized that number word inversion should also influence numerical processing *within* one and the same language. To test this assumption, we investigated English- and Polish-speaking participants, because in the otherwise consistent English and Polish number word systems, teen numbers are inverted (i.e., in English: numbers 13 to 19, in Polish: numbers 11 to 19). Employing an addition verification task, we investigated performance differences between stimulus categories with a varying number of inverted components. We were also interested in how those components affect the carry effect, a specific indicator of place-value processing in multi-digit addition. Across both Experiments, we expected (1) a significant carry effect, (2) an effect of stimulus category (i.e., increasing RTs the more summands within a problem are inverted), and (3) differential carry effects depending on the number of inverted summands (e.g., a less pronounced carry effect the more inverted summands in the respective no-carry stimulus category).

Results observed in both studies and for both non-carry and carry-problems indicated that the presence of summands reflected by inverted number words influenced the addition process. Regarding non-carry stimulus categories, problems within the *twenty* and *thirty* number range were responded to faster when only non-inverted summands were part of the problems as opposed to problems with one or two inverted summands. With regard to carry problems, the *thirty* condition allowed us to compare two different carry stimulus categories (*carry N+N* and *carry I+I;* N: non-inverted, I: inverted). Paralleling results for non-carry problems, carry problems involving only non-inverted summands were responded to faster than problems with two inverted summands. Thus, generalizing previous studies investigating inversion-related effects in addition problem solving between language groups (e.g., Göbel et al. 2014, Lonnemann & Yan, 2015), the present results indicate that number word inversion also influences addition problem solving within one and the same language group.

Furthermore, comparable to previous studies investigating processing costs resulting from a carry operation (e.g., Klein et al., 2010; Moeller, Klein, & Nuerk, 2011), we observed significant carry effects in response times for *teens*, *twens*, and *thirties* in both studies. For *twens* and *thirties* the carry effect was larger when no-carry problems involved only non-inverted summands compared to problems with one or two inverted summands. Thereby, our data suggest that two aspects influenced response time in addition problems: the need for a carry operation and inversion the summands of the addition problem. Moreover, the carry effect was argued to be an indicator of place-value processing in multi-digit addition because it specifically reflects costs related to the manipulation of place-value stacks in a carry problem (Nuerk et al., 2015). Thus, results for *twens* and *thirties* show that not only general processing speed but also place-value processing of the respective digits was affected by the inversion of number words.

However, we observed the opposite effect for the *teens* condition in both experiments. While this result was unexpected in the English-speaking sample, it was even more surprising to find it again in the Polish-speaking sample. Our initial speculation was that the reversed pattern in the English sample may be explained by the fact that the N+N category included the numbers 11 and 12. In English, as in many other languages (e.g., German: elf, zwölf), these two number words are not only non-inverted but they are also special cases. In contrast to all other two-digit number words they do not reference units and tens explicitly. As a consequence, they lack any explicit place-value information and, thus, might be processed holistically and more similarly to single-digit numbers. As such, we assumed that some place-value information coded in a number word (even if the number word is inverted) might be better for efficient place-value processing than number words with no place-value information at all.

However, in Polish all teen number words are inverted, including 11 and 12. When we rerun our analyses in the Polish data set separately for items including 11 and 12 versus items with teens larger 12, we observed the same pattern in Polish as we saw for English. These findings are incompatible with above explanation. Rather, results suggest that, independent of their actual number word form, 11 and 12 might be special cases. A possible language-invariant explanation might be related to the frequency with which certain numbers are used in written or oral language. In a cross-linguistic corpus analysis Dehaene and Mehler (1992) showed that in addition to a general decrease of frequency with numerical magnitude, some numbers – such as number 12 – show local increases in frequency. An increase in frequency of the number 12 is fairly intuitive when we consider that in many countries eggs or other consumables are sold by (half a) dozen.

Number words for 11 and 12 also seem to hold a special role when it comes to the development of counting skills in the teen number range. Miller and colleagues (1995) compared early counting skills in English- and Chinese-speaking pre-schoolers. In contrast to English, teen numbers are highly regular in Chinese (i.e., they have a non-inverted base-ten structure: shí yī (ten one), shí èr (ten two), shí sān (ten three) etc.). While no differences in counting skills for numbers up to 10 were observed, the authors found language differences favouring the Chinese-speaking children for the teen number range. When looking at individual teen numbers more closely, a clear difference in counting skills only started beyond the number 12 (cf., Figure 2 in Miller et al., 1995). Thus, specificities in the developmental trajectories in learning number words seem to substantiate that number words for 11 and 12 are processed differently than those of other teens. They might be treated more similarly to single-digit numbers and thus might be processed holistically for problems in a small (teen) number range. Additional research is needed to provide further evidence for a special status of

11 and 12 and, if this exists, to then investigate what causes this special status. Our results seem to provide initial evidence that there is something special about 11 and 12 independent of their number word form.

In contrast to between-group studies of linguistic effects on number processing, our results cannot be explained by cultural differences between groups (e.g., educational systems, (mathematics) teaching practices; Ngan Ng & Rao, 2010; Towse & Saxton, 1998), because our effects are between items within the same group. Thus, the current study belongs to a growing body of studies that try to avoid cultural confounds between groups such as studies with bilingual speakers (e.g., Prior, Katz, Mahajna, & Rubinsten, 2015), representing between-language approaches for which the cultural environment is held constant (e.g., Mark & Dowker, 2015; Colomé et al., 2010; Dowker et al., 2008; Dowker & Roberts, 2015) or studies with speakers of one language with two different number word systems (Pixner et al., 2011). However, as far as we are aware, the current study is the first within-culture, within-participant, and within-language approach to show that beyond major cultural influences, specificities of a respective number word system influence numerical processing at the item-level.

In line with previous studies on addition problem solving (Göbel et al., 2014) and other numerical tasks (e.g., number comparison: Nuerk et al., 2005; number line estimation: Helmreich et al., 2011), we observed effects of inversion in a paradigm with symbolic Arabic digits and in adult participants. This further substantiates the argument of highly automatic co-activation of number word information even in tasks for which no verbal in- or output is required. In turn, this highlights the necessity to carefully control any stimulus characteristics regarding number word formation even for tasks that only use symbolic Arabic digits and even for studies within one and the same language. Moreover, the present study focused on effects of number words in highly skilled adult participants. Effects of number word formation are usually even more pronounced in children while acquiring numerical abilities. Thus, future studies might investigate possible developmental trajectories as well as effects of varying numerical skill levels on the influence of inverted number words within one and the same language.

Conclusions

Taken together, two independent experiments in two different language groups provide first evidence for inversion-related influences on multi-digit addition *within* one number word system and show that influences of the specificities of number word structures operate on the item level. In particular, the inversion of number words with respect to the order of digits in the Arabic digit string seems to affect exactly those problems that contained inverted components (in the present case summands). As such, the present study is amongst the first to show inversion-related influences on numerical processing in a within-subject and within-language design. Thereby, the present results generalize previously observed detrimental inversion-related effects observed for addition in between-language group comparisons to the case of within-language comparisons, with only one number word system and in one and the same person. In addition, our results suggest that language-invariant factors such as the need for a carry and possibly number word frequency also contribute to multi-digit addition problem solving.

Acknowledgements

We would like to thank Dr Katarzyna Węsierska and Dr Katarzyna Sujkowska-Sobisz from the Department of Philology, University of Silesia, Poland for help with data collection and Shirley-Ann Rueschemeyer for providing the inspiration for this experiment.

Funding

This work was supported by the Leibniz-Competition Fund [SAW-2014-IWM-4].

Declaration of Conflicting Interests

The Authors declare that there is no conflict of interest.

Data Availability

The datasets generated and/or analyzed during the current study as well as the stimulus list are available in the Open Science Framework repository, https://osf.io/9jvgw/.

References

Brysbaert, M., Fias, W., & Noël, M.-P. (1998). The Whorfian hypothesis and numerical cognition: is 'twenty-four' processed in the same way as 'four-and-twenty'? *Cognition*, *66*(1), 51–77. https://doi.org/10.1016/S0010-0277(98)00006-7

Colomé, À., Laka, I., & Sebastian-Galles, 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

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

Deschuyteneer, M., De Rammelaere, S., & Fias, W. (2005). The addition of two-digit numbers: exploring carry versus no-carry problems. *Psychology Science*, *47*(1), 74–83.

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. http://doi.org/10.1080/09515080802285511

Dowker, A., & Roberts, M. (2015). Does the transparency of the counting system affect children's numerical abilities? *Frontiers in Psychology*, *6:945*. http://doi.org/10.3389/fpsyg.2015.00945

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. https://doi.org/10.3758/BF03198412

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*(1), 17–25. https://doi.org/10.1016/j.jecp.2013.10.001

Helmreich, I., Zuber, J., Pixner, S., Kaufmann, L., Nuerk, H.-C., & Moeller, K. (2011). Language effects on children's nonverbal number line estimations. *Journal of Cross-Cultural Psychology*, *42*(4), 598–613. https://doi.org/10.1177/0022022111406026

Ganayim, D., & Ibrahim, R. (2014). Number processing of A rabic and H ebrew bilinguals: Evidence supporting the distance effect. *Japanese Psychological Research*, *56*(2), 153-167. https:// doi.org/10.1111/jpr.12035

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*:*313*. https://doi.org/10.3389/fpsyg.2014.00313

Klein, E., Moeller, K., Dressel, K., Domahs, F., Wood, G., Willmes, K., & Nuerk, H.-C. (2010). To carry or not to carry – is this the question? Disentangling the carry effect in multi-digit addition. *Acta Psychologica*, 135, 67-76. http://doi.org/10.1016/j.actpsy.2010.06.002

Krinzinger, H., Koten, J. W., Horoufchin, H., Kohn, N., Arndt, D., Sahr, K., … Willmes, K. (2011). The role of finger representations and saccades for number processing: An fMRI study in children. *Frontiers in Psychology*, *2:373*. https://doi.org/10.3389/fpsyg.2011.00373

Lonnemann, J., & Yan, S. (2015). Does number word inversion affect arithmetic processes in adults? *Trends in Neuroscience and Education*, *4*(1–2), 1–5. https://doi.org/10.1016/j.tine.2015.01.002

Mark, W., & Dowker, A. (2015). Linguistic influence on mathematical development is specific rather than pervasive: Revisiting the Chinese Number Advantage in Chinese and English children. *Frontiers in Psychology, 6:203*. http://doi.org/10.3389/fpsyg.2015.00203

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. https://doi.org/10.1111/j.1467-9280.1995.tb00305.x

Moeller, K., Klein, E., & Nuerk, H.-C. (2011). Three processes underlying the carry effect in addition - Evidence from eye-tracking. *British Journal of Psychology*, 102, 623-645. http://doi.org/10.1111/j.2044-8295.2011.02034.x

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

Ngan Ng, S. S., & Rao, N. (2010). Chinese number words, culture, and mathematics learning. *Review of Educational Research*, *80*(2), 180–206. http://doi.org/10.3102/0034654310364764

Nuerk, H.- C., Moeller, K., & Willmes, K. (2015). Multi-digit number processing: Overview, conceptual clarifications, and language influences. In R. C. Kadosh & A. Dowker (Eds.), *The Oxford Handbook of Numerical Cognition* (pp. 106–139). Oxford: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199642342.013.02

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

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

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

Towse, J. N., & Saxton, M. (1998). Mathematics across national boundaries: Cultural and linguistic perspectives on numerical competence. In C. Donlan (Ed.), *The Development of Mathematics Skills* (pp. 129–150). Hove, UK: Psychology Press.

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